

**UNIVERSITÉ DU QUÉBEC**

**THÈSE**

**PRÉSENTÉE À  
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES**

**COMME EXIGENCE PARTIELLE  
DU DOCTORAT EN BIOPHYSIQUE**

**PAR**

**NORMAND BEAUDOIN**

**MÉTHODE MATHÉMATIQUE ET NUMÉRIQUE DE HAUTE PRÉCISION  
POUR LE CALCUL DES TRANSFORMÉES DE FOURIER, INTÉGRALES,  
DÉRIVÉES ET POLYNÔMES SPLINES DE TOUT ORDRE;**

**DÉCONVOLUTION PAR TRANSFORMÉE DE FOURIER ET  
SPECTROSCOPIE PHOTOACOUSTIQUE  
À RÉOLUTION TEMPORELLE.**

**Août 1999**

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

... à *Lise*,

*Élène*,

*Danielle*,

*Michèle...*

# TABLE DES MATIÈRES

TABLE DES MATIÈRES.....	3
REMERCIEMENTS .....	5
CITATIONS .....	11
DU MÊME AUTEUR .....	14
LISTE DES TABLEAUX .....	15
LISTE DES FIGURES .....	17
RÉSUMÉ .....	19
INTRODUCTION .....	21
Photorécepteurs et bactériorhodopsine .....	22
Nos objectifs .....	26
Un peu d'histoire.....	27
Position du problème .....	30
La cellule.....	31
Convolution et déconvolution .....	33
Méthodes de déconvolution .....	35
Misères des TFD ordinaires .....	38
L'omniprésence du bruit .....	44
Vérification ultime et séparation.....	49
Synoptique .....	50
PREMIER ARTICLE .....	51
Abstract .....	53
Résumé.....	54
Introduction .....	55
Fundamentals .....	58
Optimum parameters and precision .....	72
Examples .....	76
Computation time.....	83
Comparisons.....	84
Conclusion .....	88
Acknowledgments.....	89

Bibliography.....	90
DEUXIEME ARTICLE .....	92
Abstract .....	94
Résumé.....	95
Introduction .....	96
Fourier transform deconvolution .....	100
Test experiment.....	109
Experiment with purple membranes .....	116
Temperature discrimination .....	121
Conclusion .....	132
Bibliography.....	134
CONCLUSION .....	136
Synoptique .....	137
Le laser et les impulsions multiples .....	138
La cellule et le froid .....	141
La solution de référence .....	142
La puissance du laser .....	143
Logiciels, programmes et machines.....	144
Conclusion générale et perspectives .....	146
BIBLIOGRAPHIE GÉNÉRALE .....	150
ACCUSÉS DE RÉCEPTION.....	153

## REMERCIEMENTS

Il convient, dans la partie “remerciements” d’une thèse, de remercier en premier son directeur de recherches. Dans le cas présent, cependant, il ne saurait être question que de convenances.

Bien qu’il ne doive y avoir d’abîme entre recherche appliquée et fondamentale, et que l’intersection des deux soit loin d’être vide, il est cependant clair qu’il s’agit de deux facettes bien distinctes d’un même processus. En recherche appliquée, l’aspect “mise en oeuvre” prime. L’aspect “découverte” lui est subordonné. En recherche appliquée, généralement, on sait assez précisément ce que l’on doit faire. On sait d’où l’on part et où l’on veut arriver. La recherche porte surtout, bien souvent par élimination et considérant une foule de contraintes pratiques, sur la détermination des meilleurs chemins pour atteindre notre but. En recherche fondamentale, on retrouve la réciproque du paradigme. L’aspect “mise en oeuvre” sert d’étai, d’outil de preuve, de démonstration ou de vérification de la véracité de la découverte, qui, elle, doit être au premier plan. Une recherche fondamentale émane souvent d’une intuition, d’un rêve, d’une impulsion, d’un désir irrépressible de savoir, de comprendre. Le point de départ est très approximatif, le point d’arrivée est inconnu. Dans ces vastes contrées, plutôt que d’en éliminer, on y cherche tous les chemins possibles.

Trop souvent, la valeur d’un travail en recherche fondamentale est mesuré à l’aune de la recherche appliquée. Trop souvent s’immisce la tentation de diriger une recherche fondamentale comme on doit le faire pour sa plus prosaïque contrepartie. Diriger une recherche appliquée n’est certainement pas une mince tâche. Diriger une recherche fondamentale est impossible! On ne peut que l’inspirer, la nourrir, l’appuyer, la supporter... Je remercie mon directeur de recherche, M. François Boucher, pour avoir accompli cela! Il a su reconnaître le

genre de démarche dans laquelle je m'étais engagé, et il a créé les conditions favorables à son accomplissement. Sa disponibilité constante, sa compétence et, je dirais, sa puissance d'analyse des phénomènes complexes, ne sont un secret pour personne. Les quelques cours que j'ai suivi avec lui furent à la fois, pour moi simple physicien, un casse-tête et une révélation. Voyez-vous, certaines protéines comportent plus de deux particules, et il y des gens qui, quand même, y comprennent quelque chose! Pour tout cela je témoigne à M. Boucher ma plus sincère appréciation. Quand nous aurons le même âge, je lui redirai "Merci François!" Entre temps, puisse le contenu de cette thèse être en soi le plus grand remerciement que je puisse lui adresser.

Il n'y a pas de "self-made-man". Ceux qui croient en être ont encore besoin de beaucoup d'aide... Je tiens ici à remercier tous ceux et celles qui, plus ou moins dans l'ombre, ont contribué à ce travail. Je pense aux Paul Boivin, André Paquet, Jean Paquet, Gaétan Monger ainsi qu'à Alain Tessier, toujours disponibles pour moult explications, en longueur, en largeur, en profondeur mais jamais en "épaisseur". Il y a aussi le personnel de l'atelier de mécanique qui a toujours fait preuve de créativité en réalisant les montages d'une qualité allant bien souvent au delà de nos espérances. N'oublions pas les secrétaires, Denise, Denise, Rachel, Lucette et Margo dont la gentillesse et l'attention ont souvent contribué à aplanir bien des difficultés. Comment ne pas saluer au passage l'incontournable François Bellemarre qui garde tout très propre et ne nous prive jamais d'une bonne jase. Mentionnons aussi le personnel du service de l'informatique qui a toujours su dégager quelques centaines de Meg dans l'ordi central ou dégoter quelques powermacs pour nos calculs interminables. Remercions aussi tous ceux et celles du "lab en haut", qui, au fil des ans, ont maintenu les stocks de bactériorhodopsine nécessaire aux multiples manipulations. Je dois remercier aussi tout ces professeurs de mathématique, d'électronique, de physique et de biophysique qui ont contribué, probablement à grand'peine, à m'introduire quelques notions dans le cerveau. Merci!

Une atmosphère, un climat, une culture favorable sont des composantes essentielles à toute démarche de recherche. Un tel milieu propice est d'abord le fait des gens qui le constituent. Pendant le temps d'un baccalauréat en physique, d'un emploi de quelques années comme électronicien ou comme chargé de cours, et pendant le temps d'un doctorat; il me fut donné de connaître, parfois de loin, parfois de plus près, une foule de personnes. Je voudrais ici, en les nommant, les remercier pour une discussion, une plaisanterie, une information, une explication, un moment de détente ou simplement pour leur stimulante présence. Merci à : Daniel Abran, Johanne Baribeau, Diane Beaulieu, Marc Beauregard, Michèle Bernier, Nathalie Boucher, Hélène Bourque, Salah Boussad, Marc Bruellemans, Hélène Chabot, Marc Champagne, Marc Charland, Denis Charlebois, Isabelle Daigle, Norman Denicourt, André Désormeaux, Jean-Pierre Drolet, Samir Élouatik, Fouèd, Judith Gallant, Mei Mei Gao, Geneviève, Monique Gilbert, Gwenaëlle Gonnet, Michel Grandbois, Louise Guérette, Hamed, Jean Hamelin, Nathalie Hauet, Nathalie Henry, Michèle Jacob, Hassan Kassi, Kito, Nicole Lacroix, François Lamarche, Lucie Laurin, Pierre Lavigne, Thierry Lebhan, Ki-Hwan Lee, Gaétan L'Heureux, Salim Messouadi, Mario Méthot, Nathalie Miller, L. Morin, Jean-Claude Morissette, Chouhaid Nasr, E. Néonéné, Gilles Picard, Benoit Pigeon, Marc Purcell, Stefka Taneva, Abderhahim Tazi, Georges Todorov, Touffik, Michel Trudel, Lydia Vaugeois.

Je décoche une pensée particulière envers Jean-Luc Brousseau et Hugo Lavoie pour leur aide et conseils en tout ce qui touche les logiciels et les ordinateurs, mais aussi et surtout pour leur appui et leur présence qui, bien que souvent virtuels, n'en fut pas moins appréciés. Un autre tir est destiné à Sonya Grenier dont l'intelligence et la tenacité, qui l'ont menée elle aussi au terme du grand voyage, fut, tout au long du périple, un encouragement et une inspiration. Il faut cibler aussi Serge Gauvin, vieux copain des premières heures, pour sa confiance et pour toutes ces heures de discussions scientifiques allant de la physique fondamentale aux ondelettes, des tenseurs aux blindages électromagnétiques ou des auto-convolutions aux électro-dépositions sous vide. Sincèrement merci!



En même temps que mon doctorat, j'ai commencé un entraînement physique bienfaisant et nécessaire. Je remercie mon entraîneur de karaté, M. Sylvain Bertiaume, pour son professionnalisme, sa disponibilité, ses encouragements, et aussi pour son amitié. Je le remercie d'avoir gardé la "machine" en marche. Sans lui, ou bien je ne serais pas ici aujourd'hui, ou bien j'y serais... en double! Merci Sylvain!

Tenir! Voilà le précepte premier pour faire un doctorat. Quand les résultats escomptés ne sont pas au rendez-vous, quand les embûches se multiplient, quand en vertu de la crise, on a éteint la lumière au bout du tunnel; il faut tenir! Voilà la leçon que nous apprîmes ensemble, mon vieil ami Érik Harvey-Girard et moi. Merci Érik pour cette époque où, en échange de quelques équations ou circuits, tu m'apprenais, au sens propre, la vie, avec ses lipides, ses peptides grégaires, ses protéines et toute sa complexité et sa diversité. Merci pour toutes ces discussions scientifiques ou philosophiques sur à peu près tout, pour ces moments de jus d'oranges après nos combats de karaté exténuants, et pour avoir choisi, toi aussi, de prendre une rame de la grande galère. Merci d'être encore là, bien que trop souvent au bout d'une fibre optique, pour continuer, toujours, à m'apprendre la vie et... pour tenir! Merci vieux!

Que ce soit dans les rayons de bibliothèques, à la recherche du concept manquant d'une leçon de physique; dans les dédales des cours d'Expérimentale ou dans les contrées du nouveau monde informatique. Que ce soit pour mettre en selle un fortran sur ces bons vieux CDC; pour harnacher la puissance des Igor et Maple à nos concepts ou pour établir les bons chemins d'un circuit électronique. Que ce soit sur les surfaces abruptes d'un espace de paramètres où nous cherchions les optimums; en expéditions, croix de Malte optique en main, dans les profondeurs d'un laser ou au coin du feu des soirées de physique. Depuis nos premiers cours de physique jusqu'à maintenant, mon vieil ami Michel Dery et moi, avons fait le coup de feu ensemble. Merci Michel pour toutes ces heures de travail où nous cherchâmes autant les problèmes que leurs solutions, pour ces longues marches, bien souvent plus silen-

cieuses que loquace, où nous cogitions une sous-routine, un montage, une interprétation ou une équation. Merci d'avoir partagé toutes ces années avec leurs succès, leurs angoisses, leurs moments de découragements et d'espoir renaissant. Merci vieille branche!

On dit que tout se joue avant six ans. Il est bien évident qu'une formation scientifique doit commencer dans ces bas âges; de là à dire qu'à six ans tout est joué... Parfois, ça prend un peu plus de temps. Les voies de la recherche et de la découverte sont multiples et variées. Certaines sont moins longues, d'autres le sont plus; mais toutes se doivent de loger à l'enseigne de la profondeur et de la ténacité. Comme disaient mes parents, si tu veux que ça pousse longtemps, tu dois labourer profond; mais si tu laboures profond, c'est un peu plus long! Une formation scientifique commence à la première interrogation, ou mieux, à la première réponse que l'on reçoit. Elle ne se termine jamais! Pour se développer cependant, elle doit être cultivée, nourrie, encouragée. Je remercie profondément, au premier chef mes parents, et ensuite mes frères et soeurs, pour avoir fourni ces réponses, ces appuis, cette culture scientifique favorable. Un peu plus tard, suite à quelques heureux événements, s'est ajouté de beaux-parents et un tas de beaux-frères et belles-soeurs qui se sont appliqués dans le même sens. Je vous remercie tous, sincèrement, famille et belle-famille, pour votre appui inconditionnel, vos encouragements constants, votre aide multiforme et même vos bienveillantes taquineries. Merci!

Dans le vent des fleurs de l'âge, je présente, à titre posthume, à mon premier professeur de physique et de mathématiques, une pensée profonde et des remerciements sincères pour ces premières leçons. Je n'ai jamais oublié, Roger, tes cours d'algèbre, ils sont présents dans toutes mes équations! À toi, mon grand frère pour toujours!

"L'essentiel est invisible..." disait Antoine de Saint Exupéry dans Le Petit Prince. Il est un paradoxe, que les plus immatérielles choses soient les plus fortes. Les idées, les pensées,

les théories façonnent plus profondément le monde que toutes les forces brutes déchaînées. À l'origine de ces immatérialités, plus éthérés encore, il y a les rêves. Ces rêves qui depuis toujours donnent soif aux Hommes, leur font atteindre des cimes que leurs forces seules leur interdisent. Ces rêves qui constamment, sur les écueils du temps, se brisent, mais qui aussi, comme le Phénix, reviennent toujours, tenaces, parfois métamorphosés, mais indestructibles. Quand ces rêves fragiles se reconnaissent et se mêlent à d'autres, dans le torrent de la vie, ils deviennent une force véritable. Une force leur permettant de s'incarner.

Je te remercie Lise, ma compagne de toujours, pour tous tes rêves, pour tous tes engagements, pour ta profonde intelligence, pour toute ta ferveur et le feu de ta vie. Je te remercie d'avoir fondu tes rêves aux miens, d'avoir toujours été là, de toujours être là, de toujours tenir, en mers étales ou par gros temps. Je te remercie de tes appuis, de tes encouragements, de tes conseils. Je te remercie de ta patience, de tes espoirs, et, somme toute, pour ta contribution inestimable à ce travail. Peut-être un jour trouveras-tu dans les machines de ces coeurs que tu soignes, un algorithme auquel tu auras contribué. Je te remercie pour ces heures de discussion et de silence, pour ces années de temps, pour ces cafés noirs à trois heures... et pour le feu des étoiles!

Je remercie aussi de tout mon coeur ces trois rêves qui grandissent et qui maintenant me chipent mon ordi. Merci à vous Élène, Danielle et Michèle. Merci de votre patience, de votre appui, de votre compréhension. Merci pour ces aides au labo parfois, pour ces colonnes de nombres à "claviotter", ou pour ces collations qui inopinément se retrouvaient sur mon bureau. Merci pour ces rires qui nous secouent les neurones, ou pour ces taquines "enfarges" aux pieds, sous la table, que jamais, ah! jamais je ne sentais venir! Merci d'avoir été ces rayons de soleil dans les moments de tunnel noir, et ces sourires frais au petit matin, quand enfin je pouvais dire: "Ça a marché!" Merci tout simplement d'être là, de vivre, et... de commencer à rêver!

## CITATIONS

*Les causes primordiales ne nous sont point connues, mais elles sont assujetties à des lois simples et constantes que l'on peut découvrir par l'observation, et dont l'étude est l'objet de la Philosophie naturelle.*

*L'étude approfondie de la nature est la source la plus féconde des découvertes mathématiques. Non seulement cette étude, en offrant aux recherches un but déterminé, a l'avantage d'exclure les questions vagues et les calculs sans issue: elle est encore un moyen assuré de former l'Analyse elle-même, et d'en découvrir les éléments qu'il nous importe le plus de connaître, et que cette science doit toujours conserver: ces éléments fondamentaux sont ceux qui se produisent dans tous les effets naturels.*

*Les équations analytiques [...] s'étendent à tous les phénomènes généraux. Il ne peut y avoir de langage plus universel et plus simple, plus exempt d'erreurs et d'obscurités, c'est-à-dire plus digne d'exprimer les rapports invariables des êtres naturels.*

*Considéré sous ce point de vue, l'Analyse mathématique est aussi étendue que la nature elle-même; elle définit tous les rapports sensibles, mesure les temps, les espaces, les forces, les températures; cette science difficile se forme avec lenteur, mais elle conserve tous les principes qu'elle a une fois acquis; elle s'accroît et s'affermi sans cesse, au milieu de tant de variations et d'erreurs de l'esprit humain.*

*Son attribut principal est la clarté; elle n'a point de signes pour exprimer les notions confuses. Elle rapproche les phénomènes les plus divers et découvre les analogies secrètes qui les unissent. Si la matière nous échappe, comme celle de l'air et de la lumière, par son extrême ténuité, si les corps sont placés loin de nous, dans l'immensité de l'espace, si l'homme veut connaître le spectacle des cieux pour des époques successives que séparent un grand nombre de siècles, si les actions de la gravité et de la chaleur s'exercent dans l'intérieur du globe solide à des profondeurs qui seront toujours inaccessibles, l'Analyse mathématique peut encore saisir les lois de ces phénomènes. Elle nous les rend présents et mesurables, et semble être une faculté de la raison humaine destinée à suppléer à la brièveté de la vie et à l'imperfection des sens; et, ce qui est plus remarquable encore, elle suit la même marche dans l'étude de tous les phénomènes; elle les interprète par le même langage, comme pour attester l'unité et la simplicité du plan de l'univers, et rendre encore plus manifeste cet ordre immuable qui préside à toutes les causes naturelles. [...] et, si l'ordre qui s'établit dans ces phénomènes pouvait être saisi par nos sens, ils nous causeraient une impression comparable à celles des résonnances harmoniques.*

*Nous avons démontré dans cet Ouvrage tous les principes de la Théorie de la chaleur, et résolu toutes les questions fondamentales. On aurait pu les exposer sous forme plus concise, omettre les questions simples, et présenter d'abord les conséquences les plus générales; mais on a voulu montrer l'origine même de la Théorie et ses progrès successifs. Lorsque cette connaissance est acquise, et que les principes sont entièrement fixés, il est préférable d'employer immédiatement les méthodes analytiques les plus étnedues, comme nous l'avons fait dans les recherches ultérieures.*

Joseph Fourier

*Je connais la tendance de l'esprit humain à faire n'importe quoi plutôt que de penser. Certes, aucun d'entre nous n'espère réussir sans travail. Nous savons tous qu'acquérir un tant soit peu de science exige un effort intellectuel considérable, et je suis sûr que nous y sommes prêts pour avancer dans notre discipline. Mais effort intellectuel n'égale pas pensée. Et ceux qui, à grand'peine, ont acquis l'habitude de s'appliquer à leur tâche, souvent trouvent plus aisé d'apprendre une formule que de maîtriser un principe. Je m'efforcerai ici de vous montrer, et vous le vérifierez plus tard, que les principes sont fertiles en résultats, alors que les résultats seuls sont stériles. Celui qui a appris une formule est à la merci de sa mémoire, mais celui qui a maîtrisé un principe peut garder son esprit libre de formules, sachant qu'il peut en fabriquer autant que nécessaire, le moment voulu. Dois-je ajouter que, malgré le recul naturel de l'esprit devant le dur processus de la pensée, pourtant, ce processus une fois accompli, l'esprit ressent une puissance et une jouissance que l'amènent à mépriser désormais les peines et angoisses qui accompagnent son passage d'un stade de développement à un autre.*

*James C. Maxwell*

*Small minds talk about people. Average minds talk about events. Great minds discuss ideas.*

*Albert Einstein*

*The surest way to corrupt a youth is to instruct him to hold in higher regard those who think alike than those who think differently.*

*Nietzsche*

*When you find yourself on the side of the majority, it is time to reform.*

*Mark Twain*

*The trouble with the world is that the stupid are cocksure, and the intelligent are full of doubt.*

*Bertrand Russell*

*Nous sommes tous jeunes! Surtout ceux d'entre nous qui le sommes depuis longtemps!*

*Henri Lebesgue*

*La clarté, c'est de la concision à long terme.*

*M.w.A.*

*Il faut double dose d'enthousiasme pour initier de grandes oeuvres. Pour les maintenir, il faut des salaires!*

*M.w.A.*

*Le courage est gaspillé sans l'intelligence, l'intelligence est stérile sans le courage.*

*M.w.A.*

*A true measure of your worth includes all the benefits others have gained from your success.*  
Cullen Hightower

*One person with a belief is equal to a force of ninety-nine who only have interest.*  
John Stuart Mill

*The Mathematics enable us to imagine more than we can clearly think.* Freeman Dyson

*The time you enjoy wasting is not wasted time.* Bertrand Russell

*Le temps ne respecte que ce que l'on fait avec son concours!* Inconnu

*La maîtrise des principes diffère de l'acquisition des techniques. Ce n'est pas en regardant les outils de mon père que j'ai compris comment fonctionnait un moteur!* M.w. A.

$$e^{i\pi} = -1$$
 Moivre

*I hold every man a debtor to his profession, from the which as a men of course doe seeke to receive countenance and profit, so ought they of duty to endeavour themselves by way of amends, to be a help and ornament thereunto.* Francis Bacon

## DU MÊME AUTEUR

La présente thèse s'articule autour de deux articles majeurs reflétant le travail plus spécifiquement accompli pour le doctorat. Cependant, depuis le début de mes études universitaires jusqu'à la fin de ce doctorat, il s'est écoulé quelques années. Tout au long de cette période, il faut le dire au moins une fois, j'ai dû travailler à temps partiel et même à temps plein pendant plusieurs années. Ce travail m'a amené à contribuer à un certain nombre de publications autres que celles de cette thèse. Voici les coordonnées de ces autres publications:

- N.Henry, N. Beaudoin, J. Baribeau and F. Boucher, **Further Characterization of Anesthetic treated Purple Membranes**, Photochemistry and Photobiology, vol.47, No 1, pp. 85-90, (1988).
- H. Kass, S. Hotchandani, R.M. Leblanc, N. Beaudoin, M. Dery, **Pulsed photoconductivity of chlorophyll a**, SPIE-The International Society for Optical Engineering, vol. 1436, pp. 58-69 (1991).
- F. Boucher, S.G. Taneva, S. Elouatic, M. Dery, S. Messouadi, E. Harvey-Girard, N. Beaudoin, **Reversible Inhibition of Proton Release Activity and the Anesthetic-Induced Acid-Base Equilibrium between the 480 and 570 nm Forms of Bacteriorhodopsin**, Biophysical Journal, Vol. 70, pp. 948-961 (1996).
- N. Beaudoin, **Fourier Transform deconvolution of noisy signals and partial Savitzky-Golay filtering in the transformed side**, Proceedings of the 12th Conference on Vision Interface. Canada. pp. 405-409 (1999).

## LISTE DES TABLEAUX

Tableau 1	Values of $e_\theta$ (eq. (41)) for different values of $N$ and $\theta$ . Dark shaded cells indicate $\theta_{opt}$ . Notation: 1.86 over $(-09)$ stands for $1.86 \times 10^{-9}$ . Computations were performed with 200 digits. Results are for function $h$ as defined by eq. (43).	77
Tableau 2	Values of $E_\theta$ (Eq. (42)) with $R_\theta$ being the numerical first derivative of $h(t)$ , with $N = 128$ . $O$ here is a functional defined as the average on the 128 points of $ R_{\theta+2} - R_\theta $ . The theoretical error is computed in the same way but with the exact first derivative in place of $R_{\theta+2}$ . Dark shaded cell indicates $\theta_{opt}$ . Computations were performed with 50 digits. Results are for function $h$ as defined by eq. (43).	77
Tableau 3	Errors for different operations obtained with $N = 256$ and $\theta = 15$ . Computations were performed with 200 digits. Results are for function $h$ as defined by eq. (43).	78
Tableau 4	Errors on numerical derivatives for all orders up to $\theta$ . Note that even if $\theta$ must be odd, derivatives are available and accurate for even order as well. Computations were performed with 200 digits. $N = 256$ . Results are for function $h$ as defined by eq. (43).	79
Tableau 5	Connection spline condition. A.M.D. stand for absolute maximum discrepancy. Notation: 3 $(-99)$ stand for $3 \times 10^{-99}$ . Computations were performed with 100 digits. $N = 256$ . Results are for function $h$ as defined by eq. (43).	80
Tableau 6	Errors for different operations obtained with $N = 256$ and $\theta = 15$ . Computations were performed with 200 digits. Results are for function $h$ as defined by eq. (44).	81
Tableau 7	Errors on numerical derivatives for all orders up to $\theta$ . Note that even if $\theta$ must be odd, derivatives are available and accurate for even order as well. Computations were performed with 200 digits. $N = 256$ . Results are for function $h$ as defined by eq. (44).	82
Tableau 8	Value of $e_\theta$ (eq. (41)). Computations were performed with 220 and 300 digits respectively. Results are for function $h$ as defined by eq. (44).	82
Tableau 9	Relative computation time for Fourier transform (eq. (38)) with $N = 256$ . Unit time is that required to perform a FFT. Computations were performed with 25 digits. Results are for function $h$ as defined by eq. (44).	83



Tableau 10	Average absolute errors for different methods of Fourier transform applied on indicated functions. In the case of DFT+, $\theta$ represents the order of correction for a canonical- $\theta$ function. For LagPol and Present method, $\theta$ is the order of the system. $N = 256$ , $T = 1$ . Computations were performed with 100 digits.	85
Tableau 11	Average absolute errors for different methods of Fourier transform applied on function $h$ as defined by eq. (43). In the case of DFT+, $\theta$ represents the order of correction for a canonical- $\theta$ function. For LagPol and Present method, $\theta$ is the order of the system. $N = 256$ , $T = 1$ . Computations were performed with 100 digits.	86
Tableau 12	Average absolute errors for different methods of Fourier transform applied on indicated functions. $T = 1$ , $p = 6$ . For Present method, $\theta$ is the order of the system. Computations were performed with 100 digits except for results marked by * for which 150 digits were used.	87
Tableau 13	Errors in deconvolution with ordinary DFT and accurate algorithm of Fourier transform. In all cases a 4x 0-padding is used. A.M. stands for author's method. Relative RMS errors are computed relatively to exact result (eq. (49), curve 1 in Fig. 7). Relative error on peak's position is the departure of the considered peak from peak's position of the exact curve, relatively to full width at half maximum (fwhm) of curve 1.	105
Tableau 14	Comparisons between direct photodiode measurement of the exciting function $E(t)$ and its recovery by deconvolution. The letters A, B, C refer to corresponding labels in Fig. 10.	113
Tableau 15	Different results for the conformational volume change of the bacteriorhodopsin near the K to L stage of the photocycle. These changes occur in the microsecond scale (1-3 $\mu$ s).	131

## LISTE DES FIGURES


- Figure 1 Deux vues de la molécule de bactériorhodopsine. Celle de gauche est un peu plus “réaliste” mais laisse moins voir la structure globale de cette protéine ainsi que son rétinale. Celle de droite est plus schématique et on voit mieux, au centre, le rétinale, cette antenne qui capte les photons. 23
- Figure 2 Photocycle de la bactériorhodopsine. L’amplitude des changements de volume illustrés, obtenus par les différentes équipes indiquées, le sont pour une molécule. 25
- Figure 3 Comparaison entre TFD et transformée de Fourier exacte. Fonction:  $h(t) = e^{-2t}$ . Le graphique (1) montre la fonction  $h(t)$ . Le graphique (2) montre, en traits fins et noirs, les parties réelles et imaginaires de la transformée de Fourier analytique de la fonction  $h(t)$ ; en traits plus larges et gris les parties réelles et imaginaires de la DFT de la fonction  $h(t)$ , numérisée en 256 points. Les graphiques (3) et (4) montrent, respectivement, les erreurs absolues et relatives de la DFT sur les premiers 128 points; les traits noirs et fins donnant les erreurs sur la partie réelle, tandis que les traits larges et gris, sur la partie imaginaire. 39
- Figure 4 Comparaison entre TFD et transformée de Fourier exacte. Fonction:  $h(t) = e^{-12t}$ . Le graphique (1) montre la fonction  $h(t)$ . Le graphique (2) montre, en traits fins et noirs, les parties réelles et imaginaires de la transformée de Fourier analytique de la fonction  $h(t)$ ; en traits plus larges et gris les parties réelles et imaginaires de la DFT de la fonction  $h(t)$ , numérisée en 256 points. Les graphiques (3) et (4) montrent, respectivement, les erreurs absolues et relatives de la DFT sur les premiers 128 points; les traits noirs et fins donnant les erreurs sur la partie réelle, tandis que les traits larges et gris, sur la partie imaginaire. 40
- Figure 5 Comparaison entre TFD et transformée de Fourier exacte. Fonction:  $h(t) = e^{-64t}$ . Le graphique (1) montre la fonction  $h(t)$ . Le graphique (2) montre, en traits fins et noirs, les parties réelles et imaginaires de la transformée de Fourier analytique de la fonction  $h(t)$ ; en traits plus larges et gris les parties réelles et imaginaires de la DFT de la fonction  $h(t)$ , numérisée en 256 points. Les graphiques (3) et (4) montrent, respectivement, les erreurs absolues et relatives de la DFT sur les premiers 128 points; les traits noirs et fins donnant les erreurs sur la partie réelle, tandis que les traits larges et gris, sur la partie imaginaire. 41
- Figure 6 Division de fonctions quasi-identiques, ne différant que par du bruit. Le graphique (1) est la fonction  $(1 - 2x + v(x))$  où  $v$  est une fonction donnant un bruit Gaussien d’écart-type 0.01. Les graphiques (2), (3) et (4) sont des divisions de fonctions comme celle du graphique (1) 45
- Figure 7 Deconvolution of  $D(t)$  (eq. (50)) by  $I(t)$  (eq. (48)). Curve 1 shows the exact result,  $E(t)$  (eq. (49)). Curve 2, indistinguishable from curve 1, is the result obtained with the author’s method. Curve 3 and 4 are obtained with DFT using different numbers of points. 103

Figure 8	Deconvolution of noisy signals. The light shaded background is the stacking of the 64 deconvolutions of the noisy signals. The two, almost identical thin black curves are, for one the average of 57 of the 64 individual deconvolutions (AD), and the other is the deconvolution of the averages (DA) of the 64 initial noisy functions $I(t)$ and $D(t)$ respectively. The dark shaded thick curve is the theoretical noiseless result given by eq. (49)	108
Figure 9	Schematics of the time-resolved photoacoustic spectroscopy system.	109
Figure 10	Typical signals of $\int aE(t)$ measured from the photodiode (upper graph) and $\int E(t)$ obtained by deconvolution (lower graph). Parameter $a$ has been adjusted to make the amplitudes comparable. In the lower part, the thin black curves show the results of the curve fitting of the error function to the steps.	112
Figure 11	Upper graph: Amplitude profiles of the source of acoustic wave through the cell. Lower graph: Deconvolution result for two signals with non-uniform spatial distribution of the amplitude of their sources of acoustic wave.	115
Figure 12	First graph: signal from the reference compound at 11.5°C. Second graph: signal from active sample at 11.5°C. Third graph: signal from active sample at 22.5°. Each signal is the average of 32 raw signals produced by 32 laser pulses. All graphs are at the same scale.	117
Figure 13	First graph: average of deconvolutions of reference signals by reference signals at 11.5°C. Second graph: average of deconvolutions of purple membranes signals by reference signals at 11.5°C. Third graph: average of deconvolutions of purple membranes signals by reference signals at 22.5°C. All graphs are at the same scale.	119
Figure 14	Upper graph: Results of temperature discrimination of components of signals issued from deconvolutions. Curve 1 is the thermal expansion of water in the active sample. Curve 2 is the thermal expansion of water in the reference compound solution. Curve 3 is the conformational volume changes of water in the reference compound solution. Curve 4 is the conformational volume changes of br in the active solution. Lower graph: the black curve is the curve 4 of upper graph. The gray background is the stacking of all the functions which gave, when averaged, the black curve.	128
Figure 15	Conformational volume changes of br, arbitrarily corrected.	129


## RÉSUMÉ

À partir d'outils mathématiques simples et bien connus, une méthode facile à comprendre, quoique puissante, a été développée. Elle donne plusieurs résultats utiles à propos des fonctions numériques. Avec les polynômes de Taylor, les fonctions delta de Dirac et la transformée de Fourier avec son homologue discret, la TFD (DFT en anglais), nous obtenons d'une fonction numérisée, son intégrale entre n'importe quelles limites, sa transformée de Fourier sans limites de bandes et ses dérivées de tout ordre. Cette même méthode, intrinsèquement, produit les polynômes splines de tout ordre et génère automatiquement les meilleures conditions frontières possibles. Diverses procédures pour déterminer les paramètres optimums sont présentées. La structure même de la méthode permet une très bonne estimation de l'erreur sur tout résultat obtenu. Des tests effectués sur des fonctions non triviales montrent que les erreurs relatives autant qu'absolues peuvent être beaucoup plus petites que  $10^{-100}$ , et il n'y a aucune indication que de meilleurs résultats ne puissent être obtenus. La méthode fonctionne aussi bien avec des fonctions réelles que complexes; elle peut donc servir aussi pour les transformées de Fourier inverses. Sa mise en oeuvre est assez facile, particulièrement si on utilise des logiciels de calcul symbolique. Les formules, une fois établies, peuvent être efficacement implantées dans un programme compilé et rapide. La méthode demeure rapide; des comparaisons entre les temps de calcul des TFR (FFT en anglais) et des transformées calculées à différents ordres sont présentées. La précision croît exponentiellement avec l'ordre cependant que le temps de calcul augmente quadratiquement. Donc, tant que l'on pourra se le permettre, augmenter l'ordre sera bénéfique. Par exemple, pour un ordre 5, le temps de calcul est seulement 10 fois celui de la TFR, mais la précision est  $10^8$  fois meilleure. Des comparaisons avec d'autres méthodes sont présentées.

Ensuite, la pertinence d'utiliser les transformées de Fourier pour déconvoluer des signaux numériques réels et bruyants, en général, et en particulier, ceux provenant des systèmes de spectroscopie photoacoustique résolus en temps (SPART), est discutée. Les conditions nécessaires pour obtenir des résultats corrects lors de déconvolutions par transformées de Fourier sont discutées. Des simulations faites à partir de fonctions analytiques connues, avec et sans bruits, et des tests à partir de signaux expérimentaux réels et, par ailleurs, déterminés, ont montré qu'avec une méthode de déconvolution appropriée, il est possible d'obtenir les bons résultats, tant en amplitude qu'en phase. De plus, des tests révèlent que peuvent être levées les deux limitations reliées au temps lors d'expériences de SPART, soient la plus fine résolution en temps dépendant du diamètre du faisceau laser, et la durée maximale de mesure, limitée par les réflexions. Un autre problème des systèmes de SPART, non-relié aux déconvolutions mais mis en évidence par elles, et dû à la dissymétrie entre les sources des signaux acoustiques des solutions photosensibles et de référence, est aussi discuté. D'autres tests faits avec la bactériorhodopsine des membranes pourpres ont montré qu'il est possible, avec des déconvolutions par transformée de Fourier, d'obtenir des résultats visibles et plausibles. Les équations pour la discrimination en température, entre les changements de volume thermiques et conformationnels, furent établies et appliquées aux résultats des déconvolutions, pour obtenir les changements de volume qui accompagnent les changements conformationnels de la photochimie primaire de la bactériorhodopsine.



Étudiant: Normand Beaudoin



Directeur: François Boucher

## **INTRODUCTION**

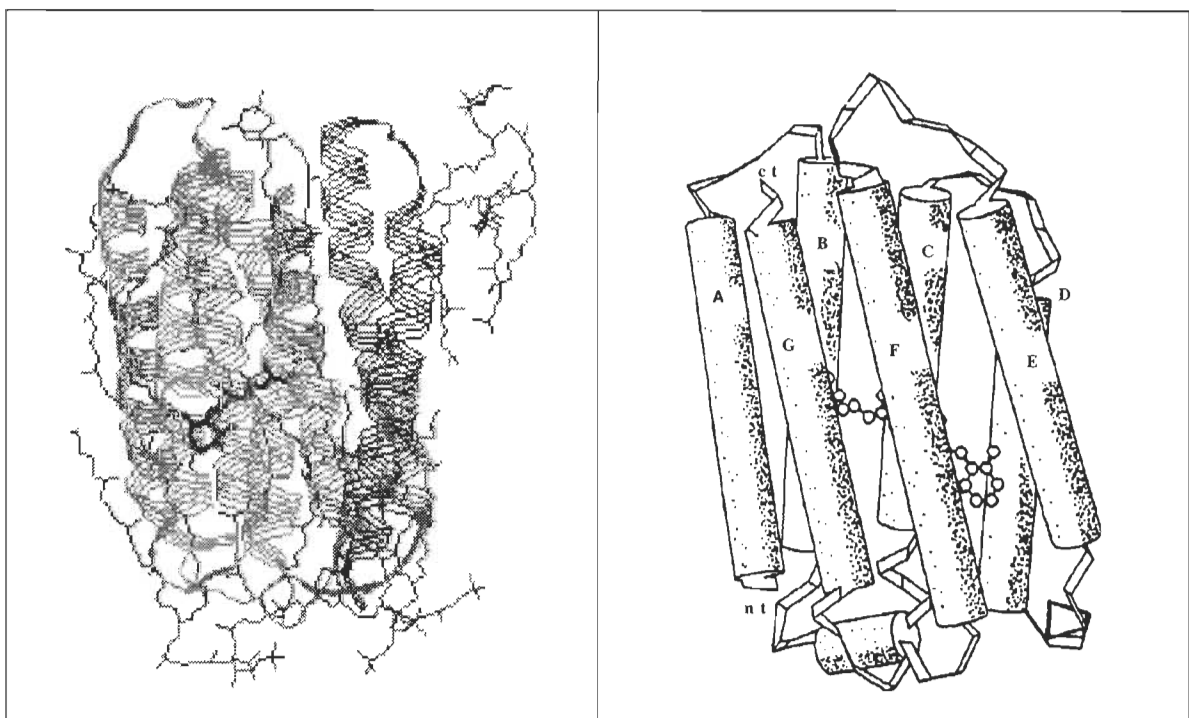
## Photorécepteurs et bactériorhodopsine

Dans les complexes photorécepteurs d'origine biologique, les mécanismes moléculaires qui dictent l'efficacité des processus de piégeage et de transduction de l'énergie électromagnétique en énergie chimique dépendent largement de l'architecture moléculaire locale ainsi que du type d'interactions qui existent entre le chromophore, responsable de l'absorption photonique, et la protéine à laquelle il est associé. Ainsi, dans les pièges chlorophylliens, les chromophores sont maintenus dans une géométrie fixe, sans liaison covalente, à l'intérieur d'un complexe protéique. L'énergie du photon absorbé est directement stockée sous forme électrostatique lorsque, dans l'état excité, le donneur primaire (une paire de molécules de chlorophylle ou de bactériochlorophylle) se voit arracher un électron par un accepteur intermédiaire (une phéophytine ou bactériophéophytine) qui le transfère ensuite à l'accepteur primaire dont le pouvoir réducteur ainsi acquis est canalisé dans une séquence de réactions d'oxydoréduction qui dissipent l'énergie initialement stockée [46]. Dans les pigments rétiens, la stratégie développée par la nature est très différente. D'abord, les pigments sont constitués par une protéine plus simple, invariablement constituée de sept hélices transmembranaires. Ces protéines ne contiennent qu'un seul chromophore, une molécule de rétinol, attachée à une lysine du squelette polypeptidique par un lien aldimine ( $-C=N+H-$ ) protoné dont la charge est stabilisée par un groupe de contre-ions fournis par des résidus d'acides aminés voisins [47].

Curieusement, le plus vieux système photosynthétique connu utilise un pigment rétinien comme piège d'énergie photonique: c'est la bactériorhodopsine. La figure 1 montre deux vues de cette molécule. Cette protéine des archéobactéries du genre *halobacterium* utilise un pigment rétinien, la bactériorhodopsine, pour convertir l'énergie solaire directement en gradient électrochimique de protons. Dans cette protéine, l'énergie d'excitation est canalisée en changements conformationnels qui modifient localement le  $pK_a$  d'acides aminés stra-

tégiques. Le résultat de ces modifications est un transport net de protons du côté cytoplasmique vers l'extérieur de la membrane. La différence de potentiel chimique des protons ainsi générée est utilisée par la bactérie pour combler ses besoins énergétiques, selon les principes de couplage standards (ATPsynthase). Ainsi, on peut qualifier la bactériorhodopsine de système photosynthétique parce que, tout comme les pigments chlorophylliens, elle dissipe l'énergie d'excitation en gradient électrochimique de protons. Il demeure néanmoins une différence importante entre la bactériorhodopsine et les systèmes chlorophylliens: chez ces derniers, le transport des protons est achevé par des chaînes oxydoréductrices qui, en plus de pomper des protons, utilisent leurs potentiels oxydant et réducteur pour produire de l'oxygène (chez les plantes supérieures) et de fixer du dioxyde de carbone.

**Fig. 1.** Deux vues de la molécule de bactériorhodopsine. Celle de gauche est un peu plus "réaliste" mais laisse moins voir la structure globale de cette protéine ainsi que son rétinal. Celle de droite est plus schématique et on voit mieux, au centre, le rétinal, cette antenne qui capte les photons.

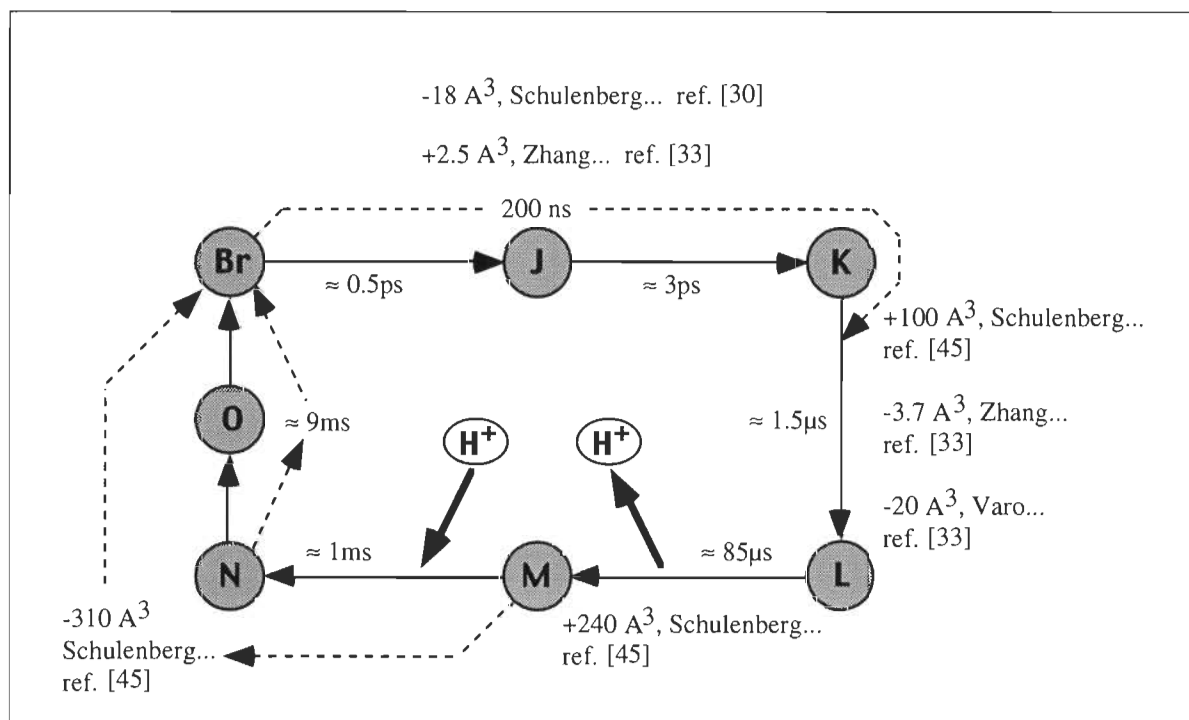




Le mécanisme moléculaire par lequel la bactériorhodopsine réussit à stocker et canaliser efficacement l'énergie du photon qu'elle absorbe n'est connu qu'en partie. En effet, en dépit du fait qu'on sache depuis longtemps que la réaction photochimique primaire de ce pigment soit l'isomérisation de son chromophore de la forme tout-trans rétinale à la forme 11-cis rétinale, que cette isomérisation survienne durant le temps de vie de l'état excité du rétinale ( $\sim 500$  fs; [48]) et que le rendement quantique de cette réaction soit élevé ( $\phi = 0,6$ ; [27]), il faut s'en remettre à des hypothèses pour expliquer comment l'énergie est transférée à la protéine. On peut imaginer, d'une part, que la relaxation du rétinale dans la forme 11-cis génère une contrainte structurale importante dans la cavité protéique qui le contient. La simple relaxation de cette contrainte pourrait se transmettre par couplage mécanique ailleurs dans la protéine et conduire aux changements de  $pK_a$  responsables du pompage de protons [49]. D'autre part, sachant que le chromophore porte une charge positive à son extrémité aldimine et que cette dernière est stabilisée par des contre-ions, il devient évident qu'une partie de l'énergie peut être stockée dans la séparation des charges qui résulte de l'isomérisation [50]. Toutefois, la quantité d'énergie qui peut être emmagasinée au cours de ce processus dépend de façon linéaire de la constante diélectrique locale et cette valeur nous échappe encore. Néanmoins, une modification locale importante des propriétés électrostatiques locales pourrait expliquer l'amorce du processus d'expulsion d'un proton. Enfin, il est également possible que le couplage entre les formes électromagnétique et conformationnelle de l'énergie soit de nature piézoélectrique. En effet, même si localement la déformation de la cavité du rétinale est minime, elle peut rompre l'organisation des liens hydrogène dans l'hélice transmembranaire qui porte la lysine 216 (celle à laquelle est attaché le rétinale dans la bactériorhodopsine) et ainsi transmettre, par couplage piézoélectrique, une charge à une distance considérable [51]. Ici encore, l'apparition d'une charge au voisinage d'un acide aminé déprotonable pourrait expliquer l'amorce du pompage de protons.

Quelle que soit l'éventualité envisagée, il est vraisemblable que l'activité de la bactériorhodopsine soit accompagnée de changements conformationnels significatifs et peut-être importants. En effet, les indices qui favorisent un mécanisme où la redistribution locale des charges produirait les changements de structure tertiaire nécessaires pour le pompage des protons commencent déjà à s'accumuler [52]. Or ce qui nous est le mieux connu sur l'activité de cette protéine, ce sont les formes spectrales par lesquelles passe son chromophore au cours d'un cycle de pompage de proton. Ce photocycle est représenté à la figure 2.

**Fig. 2.** Photocycle de la bactériorhodopsine. L'amplitude des changements de volume illustrés, obtenus par les différentes équipes mentionnées, le sont pour une molécule.



Ces changements du spectre d'absorption du pigment révèlent des modifications des interactions protéine-chromophore et cachent peut-être des modifications mesurables de la structure de la bactériorhodopsine.

## Nos objectifs

Une grande part des propriétés des protéines tiennent à leur conformation. Ainsi, la connaissance des changements conformationnels d'une protéine, les énergies mises en oeuvre lors de ces changements, les moments où ils surviennent ainsi que leur durée, revêtent une importance considérable. On peut même présumer que les variations dans ces changements, lorsqu'ils sont effectués dans des conditions différentes, soient plus révélatrices encore.

Le but initial de notre travail était d'obtenir le profil de l'ampleur de ces changements conformationnels dans le temps, particulièrement pour la bactériorhodopsine. Nous voulions carrément obtenir des graphiques dont l'abscisse est le temps et l'ordonnée, le volume. Sur ces graphiques, la courbe donnerait donc les changements de volume en  $m^3$ , en fonction du temps. Plusieurs équipes de chercheurs dans différents pays ont travaillé à obtenir ce genre de résultats avec différentes substances. Certaines sont parvenues, bien que toujours de façon indirecte, à des résultats s'approchant de notre objectif. Cependant, jusqu'à ce jour, nous n'avons vu dans aucune publication, le genre de graphique dont nous parlions plus haut. La raison est simple: c'est que les problèmes à résoudre pour y arriver ne sont pas si simples et qu'ils ne relèvent pas d'une seule discipline.

Nous avons effectivement, comme nous le verrons au deuxième article, atteint notre objectif initial. Cependant, pour ce faire, il a fallu résoudre des problèmes clé. Les solutions de ces problèmes sont devenues des objectifs prioritaires qui ont quelque peu occultés notre objectif initial, sans toutefois l'écarter. La mise au point d'une transformée de Fourier numérique précise fut un de ces importants objectifs. On peut mentionner aussi le développement de tout l'appareillage, des programmes, ainsi que les méthodes de traitement des signaux.

## Un peu d'histoire

À l'époque du début de notre travail, l'idée de mesurer ce que nous appelions alors "les tremblements de protéines" était très peu développée et la littérature sur le sujet était pour ainsi dire pratiquement inexistante. Les systèmes de spectroscopie photoacoustique (PAS) régnaient alors en maîtres dans ce domaine de la photogénération de signaux acoustiques et de leur détection. Des résultats intéressants furent ainsi obtenus concernant, par exemple, les enthalpies de réaction et/ou les énergies stockées [53, 54]. Cependant, ces systèmes, fonctionnant en fréquence et avec leur amplificateur à verrouillage de phase, ne sont pas conçus pour livrer de l'information dans le domaine du temps, si ce n'est qu'indirectement et globalement. L'usage d'impulsions lasers brèves au temps zéro, comme source actinique, au lieu d'un faisceau modulé et continu, et de détecteurs piézo-électriques, plus rapides que des microphones, devait commencer à ouvrir une nouvelle voie. Cependant, à cause de la complexité des signaux générés, ces systèmes n'utilisaient que l'amplitude du signal comme mesure. Ce n'est que très indirectement que l'information obtenue pouvait être située dans le temps. En exagérant à peine, on peut presque dire que l'on mettait à profit les défauts du système pour pouvoir dire, par exemple, que tel changement de volume devait survenir à l'intérieur de tel délai puisque c'était la limite de résolution en temps du dit système [26, 28, 29, 30, 31, 34, 35].

Pour lever ces difficultés, deux voies furent explorées: la déconvolution des signaux bruts d'une part, et la prise de signal par voie optique plutôt qu'acoustique, d'autre part. Certains auteurs ont appliqué une déconvolution aux signaux acoustiques bruts, et là encore, des résultats notables furent obtenus. Mais, comme nous en parlerons abondamment au deuxième article, à cause de la complexité des signaux et de l'usage de déconvolutions par reconvolutions itératives, encore une fois la résolution en temps et la plage temporelle d'observation devaient souffrir de sérieuses limites. En effet la résolution en temps était limitée par le diamètre du faisceau laser<sup>1</sup> et la durée maximale d'observation était limitée par les réflexions de

l'onde acoustique. En pratique, seules les toutes premières parties des signaux bruts pouvaient être soumises à la déconvolution [19, 21, 22, 23, 25, 33]. Pour repousser ces limites il eut fallu une déconvolution directe appliquée à la totalité du signal brut. Mais cela est une autre histoire, justement celle à laquelle nous essayons, bien modestement, de contribuer par ce travail.

Sommairement, la voie dite optique consiste à faire traverser un faisceau laser continu dans, ou très près de la région excitée de l'échantillon, lequel, en conséquence, doit être suffisamment transparent. L'activation de la région excitée par le laser à impulsion peut alors créer un effet de lentille ou une déviation du faisceau laser continu. L'enregistrement de ces perturbations du faisceau de mesure en fonction du temps constitue le signal brut duquel il faut tirer les informations désirées. L'extraction de ces informations repose sur une modélisation du phénomène soumise à des ajustements de paramètres. Malheureusement cette voie ne livre toujours pas, à l'instar des autres méthodes dont nous parlions plus haut, le genre de graphique "volume vs temps" que nous désirons obtenir. Par contre, jusqu'ici, cette méthode fut, la seule à exhiber une portée en temps suffisante pour observer des composantes lentes de changements de volume. Ici encore, des résultats intéressants furent obtenus [22, 44, 45]. Cependant, les résultats obtenus par cette méthode ne recoupent pas ceux obtenus par d'autres méthodes, dans des domaines de temps similaires. Curieusement aussi, tout au moins jusqu'ici, les déconvolutions ne furent pas essayées sur les signaux obtenus de cette façon. Cela viendra sans doute!

Comme nous venons de le voir, différentes méthodes furent appliquées à différentes substances, dont la bactériorhodopsine. En ce qui concerne cette dernière, les principaux ré-

---

1. Ce point est abordé au deuxième article. L'analyse et l'explication de ce phénomène nécessite la mise en équation, basé sur des considérations théoriques et sur les propriétés thermoacoustiques du milieu, du processus de génération de l'onde acoustique. Il est en soi, un sujet d'étude et a été l'objet de plusieurs publications [20, 34, 35].

sultats obtenus sont présentés avec les références pertinentes à la figure 2. Là où les recouplements le permettent, il faut remarquer la disparité des résultats, qu'ils soient obtenus par différentes méthodes ou par des méthodes semblables mais avec différentes équipes. Cela suggère que ces méthodes n'en sont sans doute pas encore à leur pleine maturité.

C'est dans ce contexte que nous entreprîmes et développâmes notre travail. Nous allons maintenant poser les problèmes auxquels nous nous sommes attaqués et décrire les difficultés que présentaient leurs solutions.

## Position du problème

Quand une molécule de bactériorhodopsine capte un photon, elle engage le cycle photochimique. Lors de ce cycle, nous présumons qu'une partie de l'énergie du photon absorbé est dégagé de façon non-radiative. Ces dégagements non-radiatifs peuvent se traduire sous forme de relaxation thermique ou sous forme de travail de type  $p dV$  résultant de changements de volume, consécutifs à une reconfiguration de la protéine. En solution dans l'eau, toute relaxation thermique de la part de la protéine, se traduira par un échauffement local de l'eau, laquelle évidemment se dilatera. Nous avons donc deux causes différentes produisant toutes deux et de façons confondues des changements de volumes. Un changement local de volume dans un échantillon engendre une onde acoustique. C'est de cette onde acoustique que la spectroscopie photoacoustique à résolution temporelle (SPART) entend tirer les informations désirées, soient les courbes des amplitudes en fonction du temps des changements de volume thermiques et conformationnels de l'échantillon.

## La cellule

Une des premières tâches donc à effectuer, pour réaliser un système SPART, est de produire cette onde acoustique, de la détecter et de l'enregistrer. À cet effet, la mise au point d'une cellule adéquate est primordiale. Puisque l'excitation de l'échantillon est réalisée par une impulsion de lumière, venant en l'occurrence d'un laser YAG, il importe qu'aucun signal parasite, i.e. aucune onde acoustique ne soit produite par aucune autre partie de la cellule que la région excitée de l'échantillon. Une cellule adéquate doit donc être exempte de toute surface susceptible d'absorber la lumière directe, réfléchie, réfractée ou diffusée. Puisque, par exemple, la cellule doit forcément être munie d'un capteur et de son support, lesquels ne peuvent, à cause de la matière dont ils sont constitués, être totalement transparents, cette condition pose des contraintes considérables. En effet, pour obtenir le meilleur rapport signal sur bruit, le capteur doit être le plus près possible de la source de l'onde acoustique produite par le faisceau laser. Par contre, pour éviter une absorption de lumière laser par le capteur, celui-ci devrait être le plus loin possible de la source acoustique.

Idéalement, puisque la valeur du coefficient d'expansion thermique de l'eau dépend de la température, la cellule doit être thermostatée. Tout système de contrôle de la température appliqué à la cellule doit, lui aussi, respecter la condition de ne produire aucune onde acoustique parasite, soit directement par ses propres vibrations, soit par absorption de lumière provenant de l'impulsion laser.

Puisque, dans un système SPART, la source d'information est une onde acoustique, la cellule doit être protégée de toutes vibrations. Elle doit donc être, le plus possible, isolée mécaniquement de son environnement. Elle doit être "découplée" du monde extérieur ou tout au moins, liée à son entourage par des liens mécaniques les plus ténus possible. Mais paradoxalement, la cellule doit être rigide et fermement maintenue, car, pour assurer une grande repro-



ductibilité du patron de l'onde acoustique, sa disposition par rapport au faisceau laser et son orientation dans l'espace doivent être tenues rigoureusement fixes. Ce dernier point est d'autant plus important que, pour des raisons explicitées dans le deuxième article, la cellule doit être vidée et remplie entre chaque mesure; ce qui nécessite une manipulation susceptible de perturber sa disposition spatiale.

Finalement, il est grandement souhaitable que le contenu de la cellule puisse être inspecté visuellement pour, par exemple, déceler la présence de bulles dans la solution, de buée sur les parois ou de tous autres éléments qui pourraient modifier l'impédance acoustique de la cellule en quelque point que ce soit.

La conjugaison de ces contraintes conduit forcément à des compromis. Nous avons conçu, dessiné, fabriqué, fait fabriquer, testé et caractérisé plusieurs types de cellules. Finalement, le type de cellule retenu, décrit dans le deuxième article, n'est pas l'idéal pour tous les cas. Cependant, nous avons choisi le type de cellule et de capteur qui, malgré d'importants compromis, sont apparus les plus appropriés pour démontrer la fonctionnalité d'un système SPART à déconvolution directe par transformée de Fourier.

## Convolution et déconvolution

Nous avons dit que c'est de l'onde acoustique que la spectroscopie photoacoustique à résolution temporelle entend tirer les informations désirées. Cela doit être précisé. À la fin d'une mesure, tout ce que nous avons c'est une fonction numérisée, enregistrée dans la mémoire d'un ordinateur. C'est d'elle que nous devons tirer nos informations, soient les changements de volume de la partie d'échantillon excitée par le laser. Mais cette fonction n'est pas le reflet fidèle de ces changements de volume. Elle est en fait la réponse de tout le système à ces changements de volume. Elle reflète toutes les oscillations induites, par ces changements, et dans la solution et dans les constituants de la cellule et du capteur. Elle incorpore toutes les réflexions de l'onde acoustique et ses atténuations; elle inclut la façon dont le capteur piézo-électrique convertit ce qu'il perçoit de l'onde acoustique en signal électrique et enfin, elle comprend toutes les inflexions imposées à ce signal électrique par les amplificateurs et leur bande passante, par les câbles coaxiaux et finalement par les convertisseurs analogues-numériques. En termes techniques, nous avons un terme-capsule pour exprimer tout cela en un mot. On dira que la fonction brute que nous numérisons et enregistrons, et que nous noterons  $D(t)$ , est la convolution de la fonction "changements-de-volume", notée  $E(t)$ , par la fonction caractéristique du système, elle étant notée  $I(t)$ . La relation mathématique entre ces fonctions est donnée par l'expression suivante:

$$D(t) = \int_0^t I(\tau)E(t-\tau)d\tau = \int_0^t I(t-\tau)E(\tau)d\tau = I(t) \otimes E(t) = E(t) \otimes I(t)$$

Nous traitons ce point plus en détail au début du premier article et nous y revenons au début du deuxième.

La fonction caractéristique du système est la réponse de tout le système à une impulsion unitaire de puissance au temps zéro. Nous parlerons d'une delta de Dirac de puissance. En pratique, pour produire cette delta de Dirac de puissance dans l'échantillon, nous utiliserons une solution de référence dont les propriétés optiques et thermoacoustiques sont identiques à celles de la solution photoactive (bactériorhodopsine) mais dont la dite substance photoactive aura été remplacée par un colorant; de l'encre noire dans notre cas. Ici, une question essentielle et délicate se pose. Est-ce que le colorant utilisé convertit suffisamment rapidement l'énergie lumineuse en chaleur pour qu'une delta de Dirac de puissance lumineuse du laser se traduise en un delta de Dirac de puissance thermique dans la solution. Incidemment, cette question a été traitée antécédemment. Il a été établi que les temps de conversion de l'énergie lumineuse en énergie thermique, notamment pour l'encre noire, se situe en deça de la nanoseconde [33].

Nous avons dit aussi, en parlant de la fonction qui est la réponse du système aux changements de volume, que c'est d'elle que nous devons tirer nos informations. Ce n'est pas faux, mais cela doit être complété. En fait, il est implicite, en parlant de convolution, que nous ayons deux fonctions en cause. Nous avons la fonction "réponse-du-système-à la-fonction-changement-de-volume" et nous avons la fonction caractéristique du système. Ce sont de ces deux fonctions que nous devons extraire l'information. Nous voulons obtenir la fonction "changements-de-volume". La fonction connue "réponse-du-système-à la-fonction-changements-de-volume" est la convolution de la fonction inconnue "changements-de-volume" par la fonction caractéristique du système qui elle aussi est connue. La fonction inconnue prend part à la convolution. Tout processus permettant d'extraire la fonction inconnue de la convolution à partir des deux autres fonctions connues est ce que nous appelons une déconvolution. Plus précisément, nous parlerons de la déconvolution de la fonction "réponse-du-système-à la-fonction-changements-de-volume" par la fonction caractéristique du système. Évidemment, tout cela sera repris, formalisé et expliqué en détail dans le premier article.

## Méthodes de déconvolution

Il existe différentes méthodes de déconvolution. L'une d'elles, très utilisée, porte le nom de déconvolution par reconvolution itérative avec critère de moindre carré. Sans entrer trop loin dans les détails, on peut la décrire comme suit. On suppose connue la fonction inconnue, ou mieux, on assume que cette fonction inconnue a une forme fonctionnelle spécifiée. On remplace donc la fonction inconnue par une fonction définie dont seuls certains de ses paramètres sont inconnus. On donne une valeur initiale approximative à ces paramètres et on calcule la convolution de cette fonction arbitraire avec la caractéristique du système. Ensuite, on compare le résultat de ce calcul, de cette convolution numérique, avec la réponse réelle du système qui est en fait la convolution physique réelle de la fonction inconnue avec la fonction caractéristique du système. Cette comparaison est effectuée via le critère de moindre carré. Le "carré" dont on parle ici est ni plus ni moins que la distance entre deux fonctions, lesquelles sont en fait deux points, ou deux vecteurs, dans un espace vectoriel de fonctions. À l'étape suivante, on remplace les paramètres arbitraires initiaux par d'autres qui selon certains critères sont susceptibles de réduire cette distance, et on recalcule une autre convolution. Ce processus est réitéré jusqu'à ce que l'on ait atteint la distance minimale entre les fonctions comparées.

Cette méthode comporte des avantages certains. C'est autant une méthode de filtration que de déconvolution. Le résultat est toujours sans bruit, car il s'agit d'une fonction analytique dont nous avons ajusté les paramètres. C'est une méthode qui se prête bien au cas où les fonctions diffèrent peu et sont de forme assez simple. Dans ces cas, modéliser la fonction inconnue et en déterminer les paramètres initiaux n'est pas trop difficile. Par contre, quand les signaux sont complexes et que la fonction inconnue est vraiment inconnue et susceptible de contenir plusieurs composantes différentes, situées à différentes valeurs d'abscisse, alors, trouver la forme fonctionnelle adéquate peut s'avérer très difficile. Et même si on y arrive, il

demeure toujours la possibilité que la fonction imposée occulte certaines composantes de la fonction cherchée. Évidemment, une inspection visuelle ou autre de la différence entre la convolution physique réelle et la dernière reconvolution numérique peut réduire ce risque. Cependant, des composantes importantes de la fonction inconnue sont souvent, par le fait même de la convolution, ténues et distribuées de façon éparse dans toute la réponse du système. Il devient alors hasardeux et peut-être même outrancièrement réconfortant de se fier au résultat trop propre et trop net d'une reconvolution itérative.

Pourtant, cette méthode demeure populaire. Jusqu'ici elle est la seule à avoir été utilisée pour la déconvolution de signaux issus de systèmes SPART [19, 21-25, 32, 33]. Il faut cependant noter que dans ces cas, n'est utilisé que seule la première alternance et souvent seule la première demie de la première alternance des signaux bruts, lesquels, peuvent en comporter quelques dizaines. Dans les systèmes SPART, cela revient à couper le signal avant qu'une première réflexion de l'onde acoustique sur une quelconque paroi de la cellule ne revienne au capteur. Les déconvolutions par reconvolutions itératives donc, se prêtent bien au cas où les fonctions sont de forme assez simples. La stratégie est alors de tronquer le signal avant qu'il ne devienne complexe... Cela réduit considérablement la richesse du signal ainsi que la durée de l'observation. En effet, tout événement qui survient après la première réflexion n'est pas observé.

Avoir une méthode de déconvolution qui n'exige aucune troncature et n'impose, a priori, aucune forme fonctionnelle ou contraintes au résultat, serait peut-être, dans bien des cas, préférable. Il demeurera toujours possible, après coup de lisser ce résultat, de le filtrer ou de lui ajuster une fonction appropriée.

Une telle méthode existe; il s'agit de la déconvolution directe par transformée de Fourier. En toute rigueur, il faut dire que calculer la transformée de Fourier d'une fonction revient

aussi à lui ajuster une autre fonction, soit une somme ou une intégrale de sinusoides, par la méthode des moindres carrés. Cependant, dans le cas de la transformée de Fourier continue, le nombre de paramètres à ajuster ainsi que leur densité sont infinis. Dans le cas de la transformée de Fourier discrète (TFD), le nombre de paramètres à ajuster est aussi grand que le nombre de points que comporte la fonction numérisée. De plus, le fait que pour chacune de ces transformées, il existe une opération inverse qui reproduise exactement la fonction initiale (sauf aux points de Gibbs, dans le cas continu), montre qu'aucune information n'est perdue.

Pourquoi alors une telle méthode de déconvolution n'est pas, en général, plus populaire; et pourquoi, en particulier, n'est-elle pas utilisée pour déconvoluer les signaux issus de systèmes SPART. Il y a deux raisons majeures à cela. Mais pour décrire ces raisons, il faut d'abord approfondir quelques notions portant sur la façon dont on calcule numériquement les transformées de Fourier et sur la façon dont on effectue la déconvolution.

## Misères des TFD ordinaires

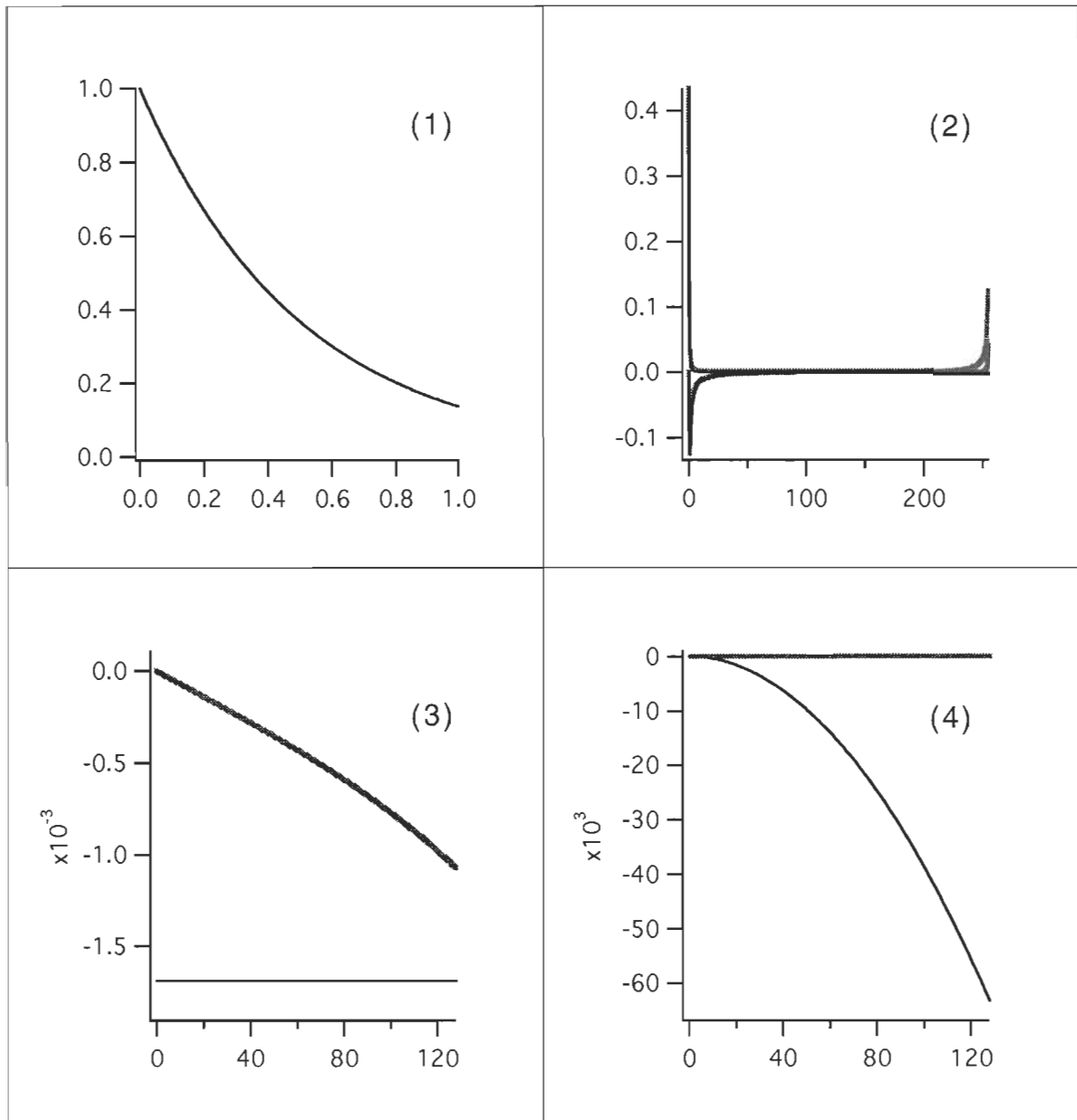
La transformée de Fourier d'un signal qui a été numérisé, est approximée par ce que nous appelons une transformée de Fourier discrète (TFD). Si la fonction initiale est numérisée en  $N$  points, alors le calcul de la TFD se fera par  $N$  intégrations numériques. Chacune de ces intégrales est effectuée avec la méthode de la sommation de l'aire de  $N$  petits rectangles. En pratique, principalement pour gagner du temps, les calculs sont effectués quelque peu différemment par des algorithmes appelés Transformée de Fourier Rapide (TFR) [Fast Fourier Transform (FFT)], mais en définitive, les résultats sont exactement les mêmes. La façon plutôt frugale d'effectuer ces intégrales est une des deux raisons pour lesquelles l'usage de la transformée de Fourier est quasi bannie des déconvolutions. C'est que cette façon de faire donne des résultats imprécis! Outre le fait que la TFD (DFT) d'une fonction est une fonction périodique alors que la transformée de Fourier d'une fonction à support borné ne l'est jamais, il arrive souvent que cette TFD, même dans sa première demi-période, soit une bien mauvaise approximation de la transformée de Fourier.

Voyons quelques exemples illustrant cette imprécision des TFD. Soit la fonction  $h(t) = e^{at}$ , pour  $t \in [0,1]$ ,  $h(t) = 0$  ailleurs. Différentes valeurs de  $a$ , uniformément réparties sur une échelle logarithmique, furent choisies pour passer de cas relativement facile pour la DFT à des cas un peu plus difficile, mais quand même très plausibles. Les figures 3, 4 et 5 montrent cette fonction, les transformées de Fourier exactes ainsi que leurs approximations données par la TFD, pour ces différentes valeurs de  $a$ . On y voit aussi les erreurs absolues et relatives de la TFD. Notons que la DFT étant périodique, nous ne pouvons utiliser, comme approximation de la transformée de Fourier, que la première moitié de ses points<sup>2</sup>. Ainsi donc, les graphiques montrant les erreurs ne portent que sur les 128 premiers points.

---

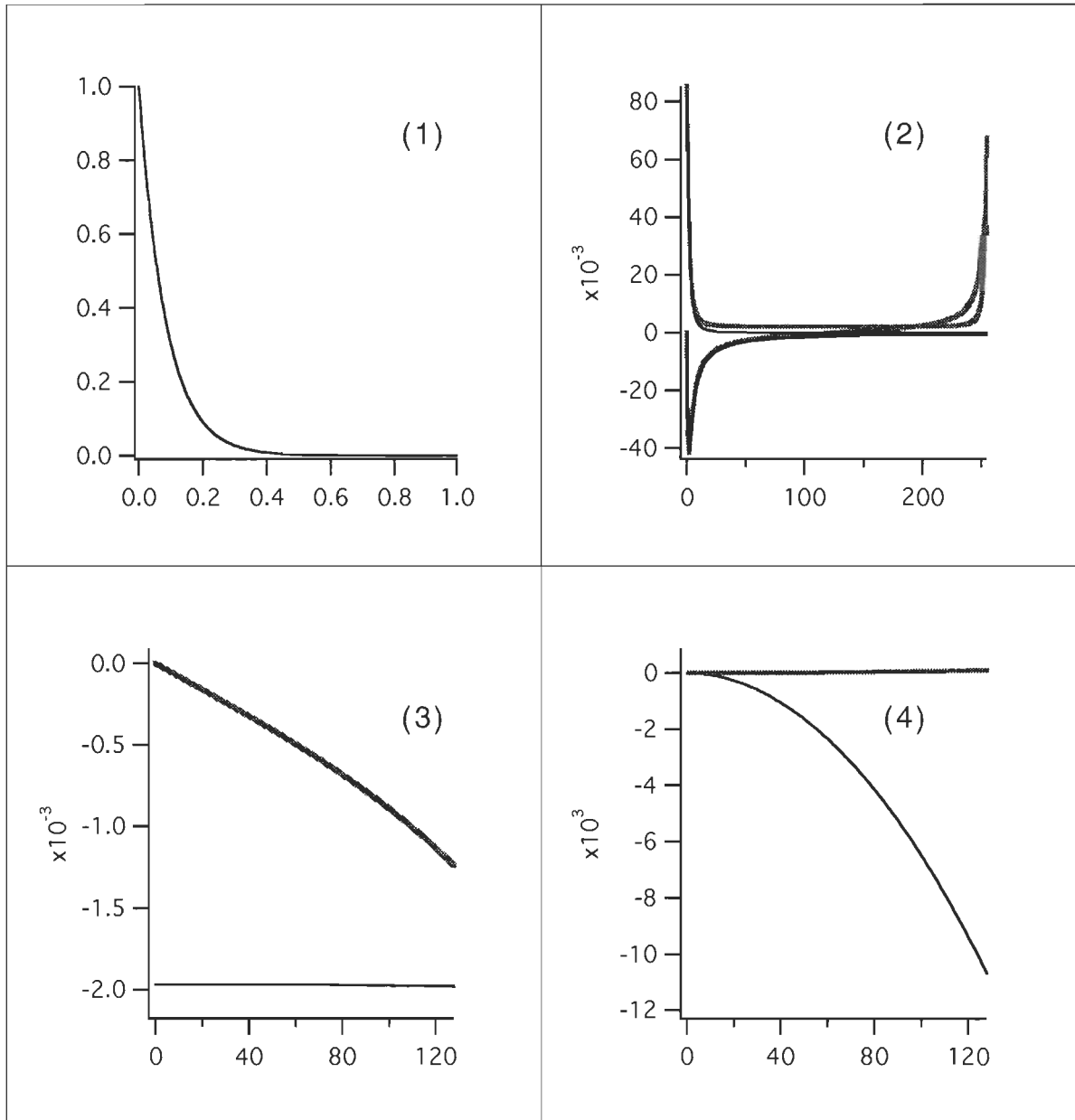
2. Cet inconvénient disparaît complètement avec la méthode présentée dans le premier article. Elle permet de calculer une excellente transformée de Fourier pour n'importe quelle valeur du spectre.

**Fig. 3.** Comparaison entre TFD et transformée de Fourier exacte. Fonction:  $h(t) = e^{-2t}$ . Le graphique (1) montre la fonction  $h(t)$ . Le graphique (2) montre, en traits fins et noirs, les parties réelles et imaginaires de la transformée de Fourier analytique de la fonction  $h(t)$ ; en traits plus larges et gris les parties réelles et imaginaires de la DFT de la fonction  $h(t)$ , numérisée en 256 points. Les graphiques (3) et (4) montrent, respectivement, les erreurs absolues et relatives de la DFT sur les premiers 128 points; les traits noirs et fins donnant les erreurs sur la partie réelle, tandis que les traits larges et gris, sur la partie imaginaire.

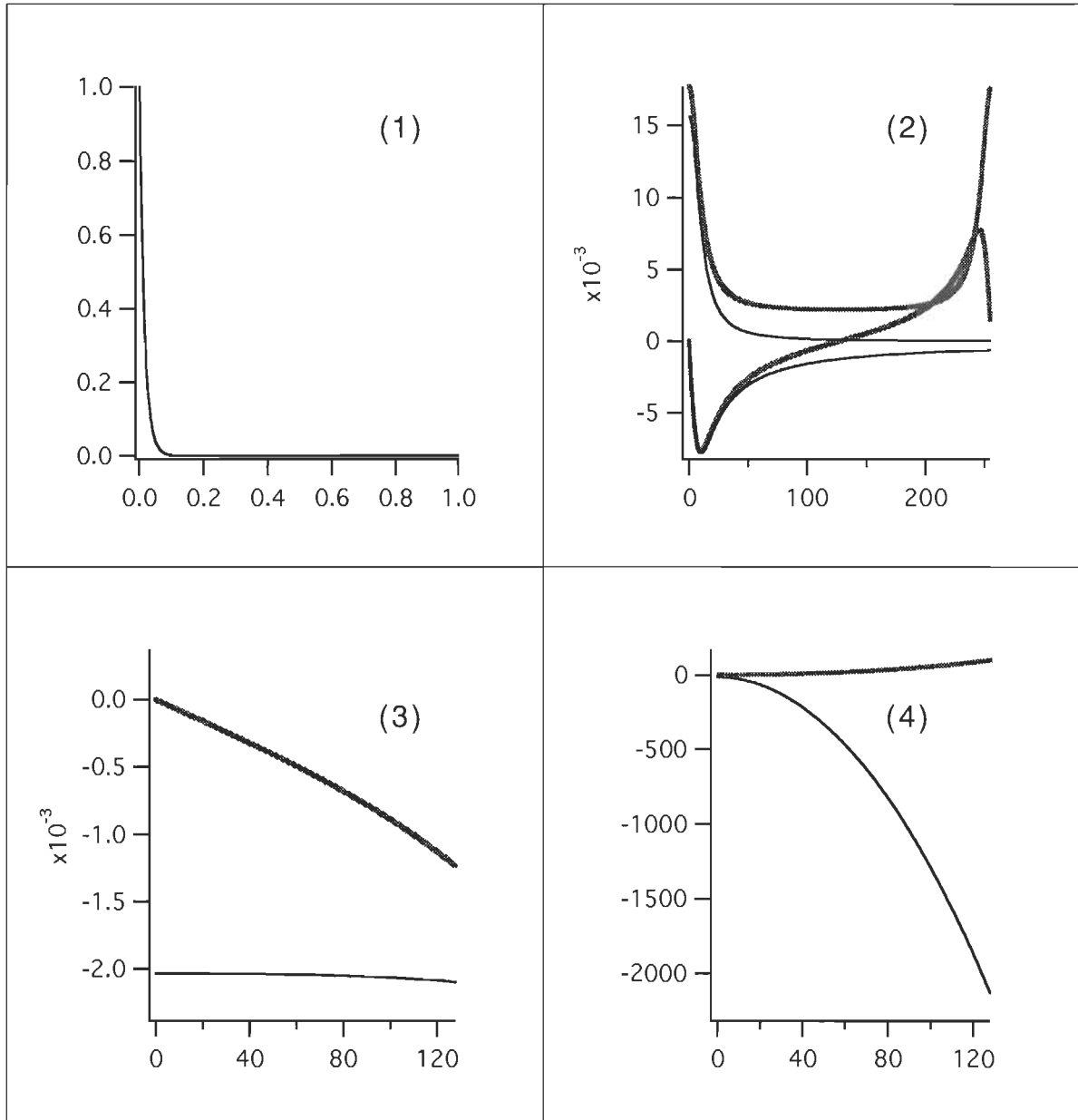




**Fig. 4.** Comparaison entre TFD et transformée de Fourier exacte. Fonction:  $h(t) = e^{-12t}$ . Le graphique (1) montre la fonction  $h(t)$ . Le graphique (2) montre, en traits fins et noirs, les parties réelles et imaginaires de la transformée de Fourier analytique de la fonction  $h(t)$ ; en traits plus larges et gris les parties réelles et imaginaires de la DFT de la fonction  $h(t)$ , numérisée en 256 points. Les graphiques (3) et (4) montrent, respectivement, les erreurs absolues et relatives de la DFT sur les premiers 128 points; les traits noirs et fins donnant les erreurs sur la partie réelle, tandis que les traits larges et gris, sur la partie imaginaire.



**Fig. 5.** Comparaison entre TFD et transformée de Fourier exacte. Fonction:  $h(t) = e^{-64t}$ . Le graphique (1) montre la fonction  $h(t)$ . Le graphique (2) montre, en traits fins et noirs, les parties réelles et imaginaires de la transformée de Fourier analytique de la fonction  $h(t)$ ; en traits plus larges et gris les parties réelles et imaginaires de la DFT de la fonction  $h(t)$ , numérisée en 256 points. Les graphiques (3) et (4) montrent, respectivement, les erreurs absolues et relatives de la DFT sur les premiers 128 points; les traits noirs et fins donnant les erreurs sur la partie réelle, tandis que les traits larges et gris, sur la partie imaginaire.



Une analyse de ces résultats graphiques montre que la TFD affiche ses moins mauvais résultats lorsque l'intégrale est facile à faire, i.e. lorsque la fonction  $h$  se rapproche d'une droite de pente nulle. Dans le cas de la figure 3, le résultat de la TFD, pour sa première moitié, est presque indistinguishable visuellement, tout au moins sur un petit graphique, de celui de la transformée de Fourier exacte. Si cette TFD est notre résultat final, qu'elle ne servira qu'à une appréciation visuelle et qu'elle ne sera destinée à aucun autre calcul, alors elle peut être acceptable. Par contre, si les résultats de la TFD sont destinés, par exemple, à des soustractions pour obtenir des différences de spectres, ou, plus critique encore, à des divisions, pour d'éventuelles déconvolutions, alors ces erreurs, même petites visuellement, peuvent devenir catastrophiques. La figure 4, et plus encore la figure 5, illustre les dangers de ne se fier qu'aux TFD, même pour des résultats qui ne seraient destinés qu'à une analyse visuelle. Bien évidemment la situation n'en devient que plus dramatique si les résultats de la TFD doivent être soumis à de subséquents calculs.

Ces imprécisions, conjuguées à un inéluctable fait mathématique et physique peut conduire à une distorsion considérable du résultat d'une déconvolution par transformée de Fourier. En effet, et nous le verrons en détail dans le premier article, une déconvolution par transformée de Fourier implique inéluctablement la division d'une fonction par une autre. Chacune de ces fonctions est une transformée d'un signal provenant du système. Il arrive que la transformée de Fourier d'un signal passe par zéro ou soit près de zéro pour de larges plages de son domaine. S'il appert que cette transformée soit celle qui, dans la division, se retrouve au dénominateur, nous avons une situation, soit de division par zéro, soit de division par des nombres très près de zéro.

En principe, quand les signaux représentent des phénomènes réels, la fonction numérateur atteindra zéro, s'il y a lieu, aux mêmes endroits que le fera la fonction dénominateur. Donc, les divisions par zéro sont des indéterminations qui peuvent être levées. Par contre les

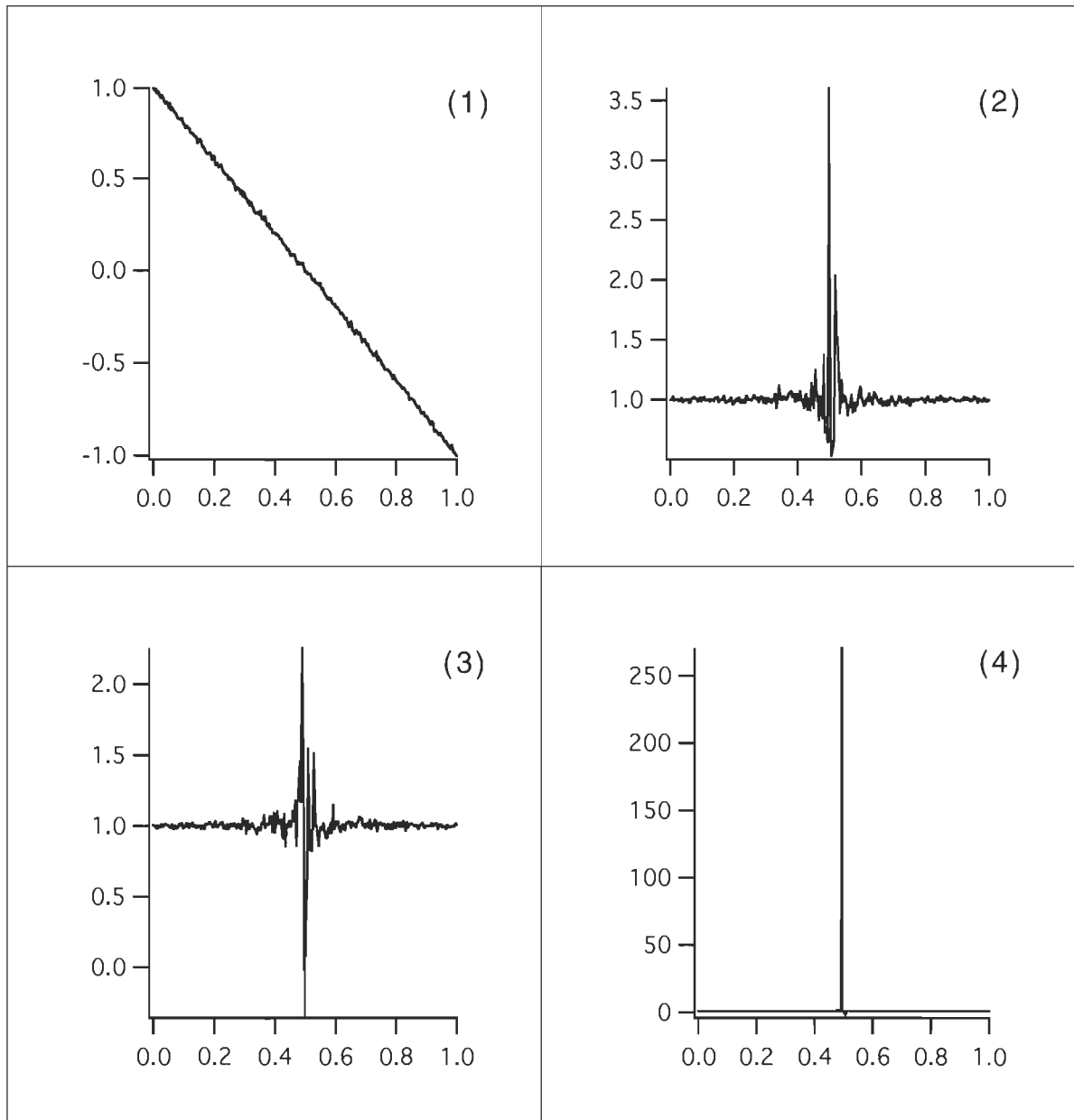
divisions par des nombres très près de zéro posent un problème. Le résultat de la division devient très sensible aux imprécisions, particulièrement à celles présentes au dénominateur. Par exemple, supposons que la méthode de calcul donnant un quelconque numérateur introduise une erreur de  $-0.001001$ . Dans le cas où nous devrions avoir  $1/1 = 1$ , nous aurions  $1/(1 - 0.001001) = 1.001002\dots$ , ce qui peut ne pas être dommageable. Mais si nous devions avoir  $0.001/0.001 = 1$  et que nous ayons  $0.001/(0.001 - 0.001001) = -1000$  alors là, c'est évidemment la catastrophe. Cet exemple porte sur des divisions en un seul point. Dans le cas de division de fonction, cela peut se reproduire plusieurs fois. Une très grande précision des transformées donc, est péremptoire; ce que n'offrent pas, à coup sûr, les TFD. La solution de ce problème occupe une part importante du premier article, lequel apporte, comme nous le verrons, des éléments intéressants sur plusieurs autres sujets.

## L'omniprésence du bruit

En plus des distorsions possibles dues aux imprécisions des TFD lors des divisions, il apparaît une autre difficulté liée, cette fois, à la nature empirique des signaux. À la différence d'une fonction mathématique pure, un signal expérimental comporte toujours du bruit. La transformée de Fourier d'un signal bruyant est une fonction comportant elle aussi du bruit. On observera dans le résultat de la division de deux fonctions bruyantes, lorsque la fonction dénominateur s'approchera suffisamment près de zéro, un phénomène qui nous apparaîtra ni plus ni moins comme une explosion du bruit. A priori, on serait tenté de dire que cette explosion de bruit perturbe, distorsionne ou même détruit le résultat de la déconvolution. Il faut être très circonspect sur ce point. Si par exemple, nous divisons une fonction bruyante  $a$  par une fonction bruyante  $b$  et obtenons comme résultat la fonction  $c$  ; il est évident qu'aux endroits où la fonction  $b$  s'approche suffisamment de zéro, nous aurons ces apparentes explosions de bruit, comme l'illustre la figure 6. Mais si on remultiplie le résultat  $c$  par la fonction dénominateur  $b$ , nous retrouverons exactement la fonction  $a$ . C'est donc que le résultat de la division est très correct. Qu'il nous plaise ou non n'a rien à voir ! Il en va de même des déconvolutions. Si nous déconvoluons une fonction bruyante  $a$  par une fonction bruyante  $b$  pour obtenir le résultat  $c$  et qu'ensuite, la reconvolution de  $c$  par  $b$  redonne bien, comme de fait, la fonction  $a$ , y inclus son bruit; nous sommes alors forcés d'admettre que le résultat  $c$ , si "bruyant" soit-il, est bien le bon.

Quand on réfléchit bien à la signification de la convolution, on peut comprendre pourquoi le résultat  $c$  doit être "bruyant" à ce point. On demande en fait au signal  $c$  de piloter un système, dont la fonction caractéristique  $b$  est donnée comme une fonction comportant du bruit, et de reproduire dans les moindres détails le signal  $a$  qui est lui aussi bruyant. Sachant

**Fig. 6.** Division de fonctions quasi-identiques, ne différant que par du bruit. Le graphique (1) est la fonction  $(1 - 2x + v(x))$  où  $v$  est une fonction donnant un bruit Gaussien d'écart-type 0.01. Les graphiques (2), (3) et (4) sont des divisions de fonctions comme celle du graphique (1).



que chaque élément de signal de  $c$  a un effet, non pas local mais global, qui se répercute, lors de la convolution, sur toute la partie subséquente de  $a$ , il devient clair que  $c$  doit être très nerveux.

Il semble y avoir une certaine “méconception” sur ce point. Il est parfois dit que les déconvolutions par transformée de Fourier sont infructueuses parce que la division génère du bruit [41, 42]. Cela nous laisse sous l'impression qu'il serait possible de trouver une façon de déconvoluer qui ne “génèrerait” pas de bruit. Une façon qui nous donnerait le même résultat qu'une déconvolution par transformée de Fourier mais sans le bruit. Cela est impossible, car ce qui est le bruit pour l'un, peut être le signal pour l'autre, et inversement. Ce qu'il faut bien saisir, c'est qu'en autant qu'on ait une densité et un nombre de points d'échantillonnage adéquat et que notre méthode de calcul de la transformée soit précise, alors, le résultat de la déconvolution par transformée de Fourier est le bon. Il n'y a pas génération de bruit!

On pourrait se demander pourquoi alors les reconvolutions itératives sont plus populaires et semblent plus fructueuses que celles par transformée de Fourier. Elles sont plus fructueuses parce qu'elles donnent le résultat que l'on veut avoir! Est-ce cela qui les rend plus populaires? Elles comportent, de façon inhérente, une grande part de filtration. Qui dit filtration dit ajouter à l'information brute des éléments d'information pré-conçus, de façon à éliminer de grands pans de cette information brute pour n'en conserver que ce que nous en désirons. La déconvolution par transformée de Fourier ne fait pas cela. Elle nous livre toute l'information. Il se peut, si nous avons une idée pré-conçue de ce qu'est le bruit et de ce qu'est le signal, que cette information, dépendamment du niveau de bruit des signaux initiaux, nous apparaisse inutilisable, physiquement inacceptable ou difficile à interpréter. Cela ne signifie pas pour autant que le résultat soit mauvais. Cela signifie que certaines composantes de ce résultat masquent certaines autres composantes que nous voudrions bien voir; c'est tout!

Il demeurera toujours possible, dans le cas d'une déconvolution par transformée de Fourier, d'injecter nos connaissances antérieures dans le processus, afin de filtrer les signaux initiaux avant de les déconvoluer et/ou de filtrer le résultat de la déconvolution. Mais, puisqu'il est toujours quelque peu périlleux de trop filtrer avant d'avoir une connaissance suffi-

sante de ce que doit être le résultat final; le fait que les étapes de déconvolution et de filtration soient nettement séparées peut contribuer à appliquer une filtration plus clairvoyante.

Maintenant, il faut battre le fer de l'autre côté! Tout ce que nous avons dit jusqu'ici ne signifie pas qu'il n'y a pas de bruit dans le résultat de la déconvolution. Il a tout simplement été dit que la déconvolution par transformée de Fourier produit un résultat conséquent et conforme aux fonctions qui lui sont soumises et qu'elle ne génère pas, par elle-même, de bruit. Cependant, si les fonctions initiales comportent une part de signal et une part de bruit, alors le résultat de la déconvolution fera de même. Il comportera une part de signal qui rendra compte de la part de signal dans les fonctions initiales, et une part de bruit qui rendra compte de la part de bruit dans ces mêmes fonctions. Ce qui peut être trompeur c'est que le rapport signal sur bruit peut ne pas être le même dans les fonctions initiales que dans le résultat de la déconvolution. Dépendamment de la structure des signaux et du bruit, si tant est que l'on puisse distinguer signal et bruit, ce rapport peut même être beaucoup plus petit dans le résultat que dans les signaux initiaux. Mais, répétons-le, il n'y a que ce qu'il faut de signal et de bruit dans le résultat final pour que, dans une éventuelle reconvolution, soient exactement reproduits les signaux et bruits de la fonction initiale.

Si toutefois, dans le résultat d'une déconvolution, le rapport signal sur bruit demeure trop faible, on peut envisager, comme nous le mentionnions plus haut, filtrer soit ce résultat, soient les signaux de départ. Le paradigme de la filtration, répétons le, consiste toujours en l'usage de connaissances antérieures pour discriminer, dans un signal brut, les composantes que nous voulons faire ressortir de celles que nous voulons éliminer. Plus on connaît les caractéristiques du signal et du bruit, plus il est possible de filtrer efficacement, mais plus cette filtration devient inutile. À la limite, une connaissance parfaite du signal et du bruit permettra une filtration parfaite... mais tout à fait inutile, puisque son objectif, déjà était atteint, avant même son application. Conséquemment, lorsqu'une filtration est utile, c'est qu'elle est im-



parfaite<sup>3</sup>. Si ces imperfections sont telles que le résultat d'une déconvolution nous apparaisse insatisfaisant, il peut être tentant de pousser plus loin la filtration. Mais alors augmentent les risques d'éliminer de l'information inconnue mais essentielle. Dans ces cas où la filtration atteint ses limites et que les résultats sont toujours trop bruyants, est-il encore possible d'obtenir un signal acceptable? Est-il possible, dans le cadre d'une déconvolution par transformée de Fourier, sans même aucune contrainte de filtration, d'obtenir un signal acceptable? La réponse est oui!

En regard de la simplicité de la méthode alternative dont nous discuterons entre autres dans le deuxième article, toute cette discussion eut apparu superfétatoire, si ce n'eut été de la nécessité, imposée par la comparaison entre les méthodes de déconvolution, de la tenir.

---

3. N'y a-t-il pas quelque chose de profond et de généralisable dans cette dernière phrase?

## Vérification ultime et séparation

Ayant un système SPART complet comprenant cellule, capteurs, amplificateur, acquisitionneur, ordinateur et méthode programmée de déconvolution par transformée de Fourier, sommes-nous enfin prêts à mesurer ces éventuels changements de volumes de la bactériorhodopsine? Pas tout à fait! Il faut nous assurer que le tout fonctionne correctement. Bien sûr, il est possible de tester des méthodes de calcul de transformées de Fourier, de déconvolution ou de réduction de bruit par des simulations numériques. Il est possible de vérifier directement la reproductibilité des signaux produits par la cellule par simple comparaison. Mais, advenant qu'un résultat final de déconvolution de signaux réels issus du système complet affiche un ou des changements de volumes, il faudrait avoir une preuve directe que ce résultat n'est pas dû à quelques propriétés intrinsèques du système physique, à un artéfact des méthodes de calcul ou à la combinaison des deux. Pour ce faire, il faudrait avoir un moyen de mimer un échantillon actif réel. Il faudrait pouvoir produire dans la cellule, à l'intérieur même de l'échantillon, des changements de volumes d'amplitudes différentes mais bien connues et ce, à différents moments décalés dans le temps. Le test ultime serait alors de voir s'il est possible, à la toute fin de la chaîne, dans les résultats même de la déconvolution, de retrouver ces mêmes changements de volume avec les bonnes amplitudes et phases. Ce point est traité en détail dans le deuxième article.

Nous avons expliqué, page 30, que les changements de volume éventuels dans la solution de bactériorhodopsine pouvaient avoir deux sources: un changement de volume dû à une reconfiguration de la protéine elle-même et/ou une dilatation de l'eau due à un dégagement de chaleur. Advenant qu'on obtienne des signaux plausibles exprimant des changements de volume dans la solution de bactériorhodopsine, il serait intéressant de pouvoir séparer ces deux composantes. Nous verrons, dans le deuxième article, ce qui peut être fait à ce sujet.

## **Synoptique**

Voilà! Nous pouvons maintenant entrer dans le vif du sujet. Le premier des deux articles qui suivent est plutôt théorique et mathématique. Sa portée est plus générale et touche un ensemble de sujets reliés à l'analyse numérique. Le second article est évidemment plus près des résultats de laboratoire. Cependant, sa première partie, traitant des conditions nécessaires pour l'obtention d'une bonne déconvolution par transformée de Fourier, est de portée plus générale et peut dépasser le cadre d'un usage pointu dans les systèmes de spectroscopie photo-acoustique à résolution temporelle. Par la suite, plusieurs éléments ne pouvant être traités dans les articles comme tel seront présentés, comme suppléments au niveau de la conclusion, juste avant la conclusion formelle.

**PREMIER ARTICLE**

**A high accuracy mathematical and numerical method  
for Fourier transform, integral, derivatives and  
polynomial splines of any order**

NORMAND BEAUDOIN,

*Groupe de Recherche sur l'Énergie et l'Information Biomoléculaire,*

*Département de Chimie-Biologie,*

*Université du Québec à Trois-Rivières,*

*3351 boulevard des Forges, C.P. 500, Trois-Rivières,*

*Québec, Canada. G9A 5H7*

(819) 376-5170

Normand\_Beaudoin@UQTR.UQuebec.ca

02.00 Mathematical Methods in Physics

02.60 Numerical approximation and analysis

Revue Canadienne de Physique / Canadian Journal of Physics

Volume 76, Numéro 9, 1998.

## Abstract

From few simple and relatively well known mathematical tools, an easily understandable, though powerful, method has been devised which gives many useful results about numerical functions. With mere Taylor expansions, Dirac delta functions and Fourier transform with its discrete counterpart, the DFT, we can obtain, from a digitized function, its integral between any limits, its Fourier transform without band limitations and its derivatives of any order. The same method intrinsically produces polynomial splines of any order and automatically generate the best possible end conditions. For a given digitized function, procedures to determine the optimum parameters of the method are presented. The way the method is structured makes easy to estimate fairly accurately the error on any obtained result. Tests conducted on non-trivial numerical functions show that relative as well as absolute errors can be much smaller than  $10^{-100}$ , and, there is no indication that even better results could not be obtained. The method works with real or complex functions as well; hence, it can be used for inverse Fourier transform too. Implementing the method is an easy task, particularly if one uses symbolic mathematical software to establish the formulas. Once formulas are worked out, they can be efficiently implemented in a fast compiled program. The method is relatively fast; comparisons between computation time for fast Fourier transform and Fourier transform computed at different orders are presented. Accuracy increases exponentially while computation time increases quadratically with the order. So, as long as one can afford it, the trade-off is beneficial. As an example, for the fifth order, computation time is only ten times greater than that of the FFT while accuracy is  $10^8$  times better. Comparisons with other methods are presented.

## Résumé

À partir d'outils mathématiques simples et bien connus, une méthode facile à comprendre, quoique puissante, a été développée. Elle donne plusieurs résultats utiles à propos des fonctions numériques. Avec les polynômes de Taylor, les fonctions delta de Dirac et la transformée de Fourier avec son homologue discret, la TFD (DFT en anglais), nous obtenons d'une fonction numérisée, son intégrale entre n'importe quelles limites, sa transformée de Fourier sans limites de bandes et ses dérivées de tout ordre. Cette même méthode, intrinsèquement, produit les polynômes splines de tout ordre et génère automatiquement les meilleures conditions frontières possibles. Diverses procédures pour déterminer les paramètres optimums sont présentées. La structure même de la méthode permet une très bonne estimation de l'erreur sur tout résultat obtenu. Des tests effectués sur des fonctions non triviales montrent que les erreurs relatives autant qu'absolues peuvent être beaucoup plus petites que  $10^{-100}$ , et il n'y a aucune indication que de meilleurs résultats ne puissent être obtenus. La méthode fonctionne aussi bien avec des fonctions réelles que complexes; elle peut donc servir aussi pour les transformées de Fourier inverses. Sa mise en oeuvre est assez facile, particulièrement si on utilise des logiciels de calcul symbolique. Les formules, une fois établies, peuvent être efficacement implantées dans un programme compilé et rapide. La méthode demeure rapide; des comparaisons entre les temps de calcul des TFR (FFT en anglais) et des transformées calculées à différents ordres sont présentées. La précision croît exponentiellement avec l'ordre cependant que le temps de calcul augmente quadratiquement. Donc, tant que l'on pourra se le permettre, augmenter l'ordre sera bénéfique. Par exemple, pour un ordre 5, le temps de calcul est seulement 10 fois celui de la TFR, mais la précision est  $10^8$  fois meilleure. Des comparaisons avec d'autres méthodes sont présentées.

## Introduction

Elaboration of the mathematical method presented in this paper has been prompted by the need for obtaining accurate Fourier transform of digitized functions in order to perform direct deconvolutions of laser induced optoacoustic signals.

The aim of deconvolution process is to retrieve information blurred by the measuring system. Suppose  $I(t)$  is the characteristic function (impulse response function) of the system and  $D(t)$  the response of that system to a causal input function  $E(t)$ . The relation between these three functions is:

$$D(t) = \int_0^t I(\tau)E(t-\tau)d\tau = \int_0^t I(t-\tau)E(\tau)d\tau = I(t) \otimes E(t) = E(t) \otimes I(t) \quad (1)$$

It is the so-called convolution of  $I(t)$  by  $E(t)$ . It is a commutative operation. When  $I(t)$  and  $E(t)$  are known, obtaining  $D(t)$  is straightforward. But in many situations, the only available signals or functions are  $I(t)$  and  $D(t)$  and what is needed is  $E(t)$ . Any process by which  $E(t)$  is obtained from  $I(t)$  and  $D(t)$  is called a deconvolution. Direct extraction of  $E(t)$  from the integrals in eq. (1) is not so easy. Fortunately, in the Fourier transformed side of the world, eq. (1) is much more simple. It is:

$$\mathcal{F}\{D(t)\} = \mathcal{F}\{I(t) \otimes E(t)\} = \mathcal{F}\{I(t)\} \mathcal{F}\{E(t)\} \quad (2)$$

A convolution in the time ( $t$ ) domain become a simple multiplication in the frequency ( $f$ ) domain.<sup>4</sup> Then, the deconvolution can be performed in the frequency domain as follow:

---

4. Convolutions or deconvolutions are not limited to causal functions and  $t$  and  $f$  can be related to others variables than time and frequency. As in [1], which is almost a classic in the field, we will use these symbols.



$$E(t) = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{D(t)\}}{\mathcal{F}\{I(t)\}} \right\} \quad (3)$$

As it is evident in eq. (3), deconvolution by Fourier transform involves a division in the frequency domain. The denominator function is the Fourier transform of the characteristic function of the system. If this denominator function is not accurate, particularly for its small values, it is evident that the deconvoluted signal could be greatly distorted.

Usually the Fourier transform of a function that has been digitized equidistantly is approximated by the so-called discrete Fourier transform (DFT). (These DFT are almost always computed via an algorithm of fast Fourier transform (FFT)). However, as it has been discussed by several authors [2-5], DFT is, in many situations, far from being adequate to act in place of Fourier transform. It is, for reasons just discussed, particularly the case for deconvolutions. Hence, it is mandatory to use numerical methods that will approximate as much as possible the behavior of exact Fourier transform. Different methods exist to do that [2-13]. In a subsequent sections, comparisons are presented for different methods and different functions.

The method developed in this paper reaches the goal of approximating, as much as one wants, the Fourier transform. Furthermore, in the process, without more complications and with the same behavior about accuracy, it produces many other results as n-derivatives, integrals, polynomial splines of any orders with their end conditions. One reason why this method remains relatively simple and can be worked out in few pages is that most computations are performed in the “Fourier transformed side of the world” where integration and differentiation become mere multiplication and division. It greatly simplifies many operations. The method performs relatively well due to its global instead of local aspect. In fact, since DFT is involved, it intermingles, in a similar way as do splines, in each of its point, information

coming from each sample point of the initial numerical function. Finally, since the whole method is based on Taylor polynomial properties, the reflection, in any of final results of those properties, can be used to estimate the accuracy of the computations.

## Fundamentals

Consider a real or complex function  $h$  of a real variable  $t$  defined as follows:

$$h(t) = \begin{cases} 0, & \text{if } -\infty < t < 0^- \\ h(t) \in \mathfrak{R} \text{ or } \mathcal{C}, & \text{if } 0^- \leq t \leq T^+ \\ 0, & \text{if } T^+ < t < \infty \end{cases} \quad (4)$$

where  $\mathfrak{R}$  and  $\mathcal{C}$  are the sets of real and complex numbers respectively, and  $0^-$  and  $T^+$  are defined as:

$$0^- = \lim_{\varepsilon \rightarrow 0} (0 - \varepsilon), \quad T^+ = \lim_{\varepsilon \rightarrow 0} (T + \varepsilon), \quad \varepsilon \in \mathfrak{R} \quad (5)$$

Furthermore, the function  $h$  is continuous and many times differentiable between  $0^-$  and  $T^+$ . At  $t = 0^-$  or  $t = T^+$ , the function can admits first kind discontinuities.

We will adopt the following definition for the Fourier transform, for  $f \in \mathfrak{R}$ :

$$\mathcal{F}\{h(t)\} = \int_{-\infty}^{\infty} h(t) e^{-i2\pi ft} dt \quad (6)$$

If we separate the interval  $[0, T]$  into  $N$  equals  $\Delta t = T/N$  parts, eq. (6), considering eq. (4), can be re-written as:

$$\mathcal{F}\{h(t)\} = \sum_{j=0}^{N-1} \left\{ \int_{j\Delta t}^{(j+1)\Delta t} h(t) e^{-i2\pi ft} dt \right\}, \quad j \in \mathfrak{N} \quad (7)$$

$\mathbb{N}$  is the set of non-negative integers. Since  $h$ , between and at 0 and  $T$ , is a well behaved function, it can be approximated, for  $t \in [j\Delta t, (j+1)\Delta t[$ , and that, for each  $j \in [0, N-1]$ , by a Taylor expansion as follows:

$$h(t) = \sum_{p=0}^{\infty} \frac{h_j^{(p)}(t-j\Delta t)^p}{p!}, \quad p \in \mathbb{N}, \quad t \in [j\Delta t, (j+1)\Delta t[ \quad (8)$$

where  $h_j^{(p)}$  is the  $p^{th}$  derivative of  $h$  at the point  $t = j\Delta t$ . Merging eq. (7) and (8) gives:

$$\mathcal{F}\{h(t)\} = \sum_{j=0}^{N-1} \left\{ \int_{j\Delta t}^{(j+1)\Delta t} \left( \sum_{p=0}^{\infty} \frac{h_j^{(p)}(t-j\Delta t)^p}{p!} \right) e^{-i2\pi f t} dt \right\} \quad (9)$$

Introducing a new variable  $\tau = t - j\Delta t$  and permuting adequately the integration and summations on  $j$  and  $p$ , eq. (9) can be written as:

$$\mathcal{F}\{h(t)\} = \sum_{p=0}^{\infty} \left\{ \left( \int_0^{\Delta t} \frac{\tau^p e^{-i2\pi f \tau} d\tau}{p!} \right) \left( \sum_{j=0}^{N-1} h_j^{(p)} e^{-i2\pi f j \Delta t} \right) \right\} \quad (10)$$

The integral in eq. (10) becomes independent of the function  $h$  or of its derivatives. It depends only on the variable  $f$  and on parameters  $p$  and  $\Delta t$ . It can be exactly evaluated analytically, once for all, for each values of  $p$ . It will be denoted as follows:

$$I_p(f) = \frac{1}{p!} \int_0^{\Delta t} \tau^p e^{-i2\pi f \tau} d\tau \quad (11)$$

The summation on  $j$  in eq. (10) is almost, as we will see later, the discrete Fourier transform of the sequence  $h_j^{(p)}$ ,  $j \in [0, N-1] \subset \mathbb{N}$ , except that here,  $f$  is continuous. We will denote it as:

$$F_p(f) = \sum_{j=0}^{N-1} h_j^{(p)} e^{-i2\pi f j \Delta t} \quad (12)$$

Substituting eq. (11) and (12) in eq. (10), we obtain the following concise equation:

$$\mathcal{F}\{h(t)\} = \sum_{p=0}^{\infty} I_p(f) F_p(f) \quad (13)$$

This last result is cardinal in that it would gives the exact values of the Fourier transform of a sampled function  $h$  (for  $j \in [0, N-1] \subset \mathbb{N}$ ), as far as we could know all the  $p^{th}$  derivatives of  $h$  for each  $j$ . According to the definition we gave at eq. (4), even if the function  $h$  is continuous and many times differentiable, its explicit analytical form is not necessarily known. It could even be known for only some of its discrete values noted  $h_j$ , as it is the case, for example, when one performs a digital measurement of a physical phenomenon. Hence, to devise a method that can be practical in these situations, we have to, and we will, develop the method considering that all  $p^{th}$  derivatives of  $h$ , for  $p > 0$ , are totally unknown. Then, it may seems that, in these conditions, eq. (13) would be completely useless. Fortunately, as we will see, we even don't have to know these derivatives, though we can, and will, obtain them as a result.

Let us then carefully examine  $h^{(n)}$ . If the function  $h$  and its derivatives are all continuous, even at  $t = 0^-$  and  $t = T^+$ , there is no ambiguities. But if there are some discontinu-

ties of the first kind at these points, we must take care of it. Since the derivative of a Heaviside unit step function [14] is a Dirac delta function<sup>5</sup>, it is well known that the derivative of a function, at a discontinuity of the first kind, is the algebraic height of the step times a Dirac delta function shifted at the point of discontinuity. The second derivative, at such a point, is handled in the same way but we should not omit to apply the differentiation process to the term containing the Dirac delta function issued from the first differentiation. The same process is successively applied for higher order derivatives. Therefore, let us denote by  $D_n$  the sum of the  $n^{th}$  derivative of  $h$  for  $t \in ]0^- - \varepsilon, 0^- + \varepsilon[$  and for  $t \in ]T^+ - \varepsilon, T^+ + \varepsilon[$ , when  $\varepsilon \rightarrow 0$ . Thus we have:

$$D_n = \begin{cases} 0, & \text{if } n = 0 \\ \sum_{m=0}^{n-1} \{h^{(m)}(0)\delta^{(n-m-1)}(t-0^-) - h^{(m)}(T)\delta^{(n-m-1)}(t-T^+)\}, & \text{if } n \in \mathbb{N} \setminus \{0\} \end{cases} \quad (14)$$

where  $\delta$  is the Dirac delta function.

The same process that has been applied to  $h$  to obtain eq. (13) can also be applied to  $h^{(n)}$ . But this time, since  $h^{(n)}$  contains delta functions and derivatives of possibly many orders of delta functions, to obtain the correct result we must add the contribution due to  $D_n$ .

Doing so we get<sup>6</sup>:

$$\mathcal{F}\{h^{(n)}(t)\} = \left( \sum_{p=0}^{\infty} I_p F_{p+n} \right) + \mathcal{F}\{D_n\} \quad (15)$$

---

5. Strictly speaking we should name it a distribution or a functional.

6. For the sake of conciseness in the formulas, the argument  $(f)$  will often be omitted for  $I_p$  and  $F_p$ . Sometimes, the same applies for  $(t)$  with functions of  $(t)$ .

On the other hand, the definition of the Fourier transform of  $h^{(n)}(t)$  is:

$$\mathcal{F}\{h^{(n)}(t)\} = \int_{-\infty}^{\infty} h^{(n)}(t) e^{-i2\pi f t} dt \quad (16)$$

Integrating by parts we get:

$$\mathcal{F}\{h^{(n+1)}\} = i2\pi f \mathcal{F}\{h^{(n)}\} \quad (17)$$

This result is valid and account for the contribution of  $D_n$  to the Fourier transform.

Defining  $b_n = i2\pi f \mathcal{F}\{D_n\} - \mathcal{F}\{D_{n+1}\}$  and combining eq. (15) and (17) gives:

$$\sum_{p=0}^{\infty} I_p F_{p+n+1} - i2\pi f \sum_{p=0}^{\infty} I_p F_{p+n} = b_n \quad (18)$$

Using the properties of Fourier transform and those of Dirac delta functions [15], one could easily see that expanding  $b_n$  results in the simple following form:

$$b_n = h^{(n)}(T) e^{-i2\pi f T} - h^{(n)}(0) \quad (19)$$

By reorganizing the terms, eq. (18) can be written as follows:

$$-i2\pi f I_0 F_n + \sum_{p=1}^{\infty} (I_{p-1} - i2\pi f I_p) F_{p+n} = b_n \quad (20)$$

As mentioned earlier, the summation on  $j$  in eq. (12) is almost the discrete Fourier transform of the sequence  $h_j^{(p)}$ ,  $j \in [0, N-1] \subset \mathbb{N}$ , except that  $f$  is continuous. In partic-

ular, if we select discrete values of  $f$  such that  $f = k\Delta f = k/T$ ,  $k \in \mathbb{N}$ , and since  $\Delta t = T/N$ , that summation will now take the exact form of the discrete Fourier transform [1] of  $h_j^{(p)}$ , which will be denoted as follows:

$$F_{p,k} = \sum_{j=0}^{N-1} h_j^{(p)} e^{-i2\pi \frac{kj}{N}} \quad (21)$$

In this context, eq. (20) will become:

$$-i2\pi f I_0 F_{n,k} + \sum_{p=1}^{\infty} (I_{p-1} - i2\pi f I_p) F_{p+n,k} = b_n \quad (22)$$

in which it is understood that  $f = k\Delta f$ . In the same way, let us define  $H_k$  as the discrete version of  $\mathcal{F}\{h(t)\}$  for  $f = k\Delta f = k/T$ , such that eq. (13) will become:

$$H_k = \sum_{p=0}^{\infty} I_p F_{p,k} \quad (23)$$

It should be noted that if  $f = k/T$ , then eq. (19) becomes a much more simple and significant expression that is:

$$b_n = h^{(n)}(T) - h^{(n)}(0) = h_N^{(n)} - h_0^{(n)} \quad (24)$$

Up to this point,  $p$  tends towards infinity. We will now introduce approximations by limiting the range of  $p$ . Let us introduce  $\theta \in \mathbb{N}$ , the truncating parameter, which, for reasons discussed later, will always be chosen as an odd integer. We will refer to it as the order of the system.



Let us expand eq. (22) for each values of  $n \in [0, \theta - 1] \subset \mathbb{N}$ . It will generate a system of  $\theta$  different equations and, for each of these, we will let  $p$  range from 1 to  $\theta - n - 1$ . This will gives the system:

$$\begin{bmatrix} (-i2\pi f I_0) & (I_0 - i2\pi f I_1) & \dots & (I_{\theta-2} - i2\pi f I_{\theta-1}) \\ 0 & (-i2\pi f I_0) & \dots & (I_{\theta-3} - i2\pi f I_{\theta-2}) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & (I_{-1} - i2\pi f I_0) \end{bmatrix} \begin{bmatrix} F_{0,k} \\ F_{1,k} \\ \dots \\ F_{\theta-1,k} \end{bmatrix} \cong \begin{bmatrix} b_0 \\ b_1 \\ \dots \\ b_{\theta-1} \end{bmatrix} \quad (25)$$

or, in a more compact format as:

$$\mathbf{M}_a \mathbf{F}_a \cong \mathbf{B} \quad (26)$$

Note that the matrix  $\mathbf{M}_a$  is completely known<sup>7</sup>. The general expression for elements of  $\mathbf{M}_a$  is:

$$(\mathbf{M}_a)_{\mu\nu} = I_{\nu-\mu-1} - i2\pi f I_{\nu-\mu} \quad (27)$$

For some combinations of possible values of  $\mu$  and  $\nu$ , the index  $p$  of  $I_p$  is negative. At that time, we must have  $I_p = 0$ . When time comes to program, the best way to handle it is to put  $I_p = 0$  when  $p < 0$ . But, from a mathematical point of view, we even don't have to do that. If we consider  $p!$  in the context of the Gamma function [16], the mere fact that  $p! = \pm\infty$  when  $p$  is a negative integer guarantees that, in these cases,  $I_p$  will always be 0, since the integral in the definition of  $I_p$  is finite.

---

7. The index  $a$  present in this matrix equation is only to distinguish its terms from those, similar but different, of subsequent matrix equations that will use index  $b$  and  $c$ .

In eq. (26), if  $\mathbf{B}$  was known, it may seem that it could be possible to compute  $\mathbf{F}_a$ . We would obtain it from the following equation:

$$\mathbf{F}_a \equiv \mathbf{M}_a^{-1} \mathbf{B} \quad (28)$$

Unfortunately, for  $f = \kappa N \Delta f$ ,  $\kappa \in \mathbb{N}$ , the solutions would irremediably diverge towards  $\pm\infty$ . However, under few conditions discussed later, the result, for values of  $f = k \Delta f$  around  $N \Delta f / 2$ , is quite accurate. We will take advantage of this fact to compute  $\mathbf{B}$ .

We must remind that the first element of  $\mathbf{F}_a$ , which is  $(\mathbf{F}_a)_1 = F_{0,k}$ , is the DFT of the sequence  $h_j$ ,  $j \in [0, N-1] \subset \mathbb{N}$ . It is completely determined for each value of  $k$ . It is not the same situation for the others elements of  $\mathbf{F}_a$  which are still unknown. Furthermore, all elements of matrix  $\mathbf{M}_a^{-1}$  are all well known for each value of  $f = k \Delta f$ . We can then extract the following from eq. (28):

$$F_{0,k} \equiv (\text{row}_1 \mathbf{M}_a^{-1}) \mathbf{B} = \sum_{\mu=1}^{\theta} (\mathbf{M}_a^{-1})_{1,\mu} (\mathbf{B})_{\mu} \quad (29)$$

Let us now define an interval of  $\theta$  values of  $k$ , centred on  $N/2$ , as follows:

$$\Omega = \left[ \frac{N}{2} - \left( \frac{\theta-1}{2} \right), \frac{N}{2} + \left( \frac{\theta-1}{2} \right) \right] = [k_1, k_2, \dots, k_{\theta}] \subset \mathbb{N} \quad (30)$$

It should now be clear why  $\theta$  is chosen as an odd integer. One should note that, at this point, everything would work fine even if  $\theta$  was not odd and if  $\Omega$  (slightly differently de-

finer) was not exactly centred on  $N/2$  but, accuracy would be slightly lower. Other stronger reasons for choosing odd values for  $\theta$  will appear later.

Let us expand eq. (29) for each value of  $k \in \Omega$ . (It is understood that in practical cases, for each instance of  $f$  in each term, one would have to replace it by  $k\Delta f$ .) Doing so gives the following system:

$$\begin{bmatrix} F_{0,k_1} \\ F_{0,k_2} \\ \dots \\ F_{0,k_\theta} \end{bmatrix} \equiv \begin{bmatrix} \left( (\mathbf{M}_a^{-1})_{1,1} \Big|_{k_1} \right) \left( (\mathbf{M}_a^{-1})_{1,2} \Big|_{k_1} \right) \dots \left( (\mathbf{M}_a^{-1})_{1,\theta} \Big|_{k_1} \right) \\ \left( (\mathbf{M}_a^{-1})_{1,1} \Big|_{k_2} \right) \left( (\mathbf{M}_a^{-1})_{1,2} \Big|_{k_2} \right) \dots \left( (\mathbf{M}_a^{-1})_{1,\theta} \Big|_{k_2} \right) \\ \dots \\ \left( (\mathbf{M}_a^{-1})_{1,1} \Big|_{k_\theta} \right) \left( (\mathbf{M}_a^{-1})_{1,2} \Big|_{k_\theta} \right) \dots \left( (\mathbf{M}_a^{-1})_{1,\theta} \Big|_{k_\theta} \right) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \dots \\ b_{\theta-1} \end{bmatrix} \quad (31)$$

Note that  $(\alpha|_{k_x})$  means the value of  $\alpha$  when  $k = k_x$  and  $f = k_x \Delta f$ . Let us express eq. (31)

in a more compact format as  $\mathbf{F}_c \equiv \mathbf{W}\mathbf{B}$ , from which we directly deduce the matrix equation that gives an important part of what we were looking for: the end conditions. It is:

$$\mathbf{B} \equiv \mathbf{W}^{-1} \mathbf{F}_c \quad (32)$$

Even if  $\mathbf{B}$  is now well known, eq. (28) cannot be used, for reasons mentioned earlier, to evaluate  $\mathbf{F}_a$ . Considering the fact previously mentioned, that the first element of  $\mathbf{F}_a$  is known, we will again expand eq. (22) but in a slightly different manner that we did to obtain eq. (25). We will once more expand it for each values of  $n \in [0, \theta - 1] \in \mathbb{N}$ . It will generate a system of  $\theta$  different equations, and for each of these we will let  $p$  range, at this time, from

1 to  $\theta - n$ . Thereafter, the terms are reorganized to obtain the following system, which will be written again in a matrix form:

$$\begin{bmatrix} (I_0 - i2\pi f I_1) & (I_1 - i2\pi f I_2) & \dots & (I_{\theta-1} - i2\pi f I_\theta) \\ (-i2\pi f I_0) & (I_0 - i2\pi f I_1) & \dots & (I_{\theta-2} - i2\pi f I_{\theta-1}) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & (I_0 - i2\pi f I_1) \end{bmatrix} \begin{bmatrix} F_{1,k} \\ F_{2,k} \\ \dots \\ F_{\theta,k} \end{bmatrix} \cong \begin{bmatrix} b_0 \\ b_1 \\ \dots \\ b_{\theta-1} \end{bmatrix} + \begin{bmatrix} i2\pi f I_0 F_{0,k} \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (33)$$

In a more compact form, we can write eq. (33) as:

$$\mathbf{M}_b \mathbf{F}_b \cong \mathbf{B} + \mathbf{C} \quad (34)$$

The general expression for elements of  $\mathbf{M}_b$  is

$$(\mathbf{M}_b)_{\mu\nu} = I_{\nu-\mu} - i2\pi f I_{\nu-\mu+1} \quad (35)$$

By comparing  $\mathbf{F}_b$  and  $\mathbf{F}_a$ , we see that the last term of  $\mathbf{F}_b$  is one order higher than the last term of  $\mathbf{F}_a$ . It corresponds to one more term in the Taylor expansion, such that the value of  $\theta$  is the same as the higher order of the derivatives in that expansion. That is why  $\theta$  is called the order of the system. Let us now re-write eq. (34) as follows:

$$\mathbf{F}_b \cong \mathbf{M}_b^{-1} (\mathbf{B} + \mathbf{C}) \quad (36)$$

The advantage of eq. (28) over eq. (36) is that it permits one to compute  $\mathbf{B}$  but, unfortunately, it shows singularities for  $f = \kappa N \Delta f$  which prevent us to compute  $\mathbf{F}_a$ . Conversely, the advantage of eq. (36) is that it exhibits a higher order than eq. (28) and it uses the information from all the points of  $F_{0,k}$ . Furthermore, provided that  $\theta$  is odd,  $\mathbf{F}_b$  is now computable for all values of  $f = k \Delta f$ . One should note that for  $f = \kappa N \Delta f$ , undetermined values

appear but, they can always be all solved by Hospital's rules, for any odd orders. For even orders, there are singularities<sup>8</sup> at  $f = (\kappa + 1/2)/(\Delta t)$ . This is one more and an imperative reason to choose odd values for  $\theta$ .

With the knowledge of  $\mathbf{F}_b$ , all the terms of eq. (23), for  $p \in [0, \theta]$ , are known. Thus, the truncated version of eq. (23) can be written as follows:

$$H_k \cong \sum_{p=0}^{\theta} I_p F_{p,k} \quad (37)$$

Let us define a one row matrix as  $\mathbf{I}_{\theta} = [I_1 \ I_2 \ \dots \ I_{\theta}]$ , and re-write eq. (37) as follows:

$$H_k \cong I_0 F_{0,k} + \mathbf{I}_{\theta} \mathbf{F}_b \quad (38)$$

With eq. (38), we have reached one of our main goals. We have a way to compute an approximation of the Fourier transform<sup>9</sup> of a digitized function. This approximation being, as we will see later, almost as much accurate as one want. This behavior is indeed related to the same behavior of Taylor expansions. Furthermore, eq. (38) is not, as is the DFT, band limited. It remains valid and accurate for values of  $k$  greater than  $N/2$  and even greater than  $N - 1$ . The DFT of a function consisting in  $N$  samples is periodic with a period  $N\Delta f$ . It is a crude approximation of the Fourier transform, only for values of  $k$  up to  $N/2$ . Beyond that point, values of DFT reappears symmetrically<sup>10</sup>, and past the point  $k = N$ , they repeat peri-

---

8. These undetermined or singular points can be moved, if the limits of integration in eq. (7) are shifted by a value  $a \in [-\Delta t, 0]$  but, they cannot be removed. The same kind of difficulties and solutions are encountered when one wants to use even order splines instead of odd order splines [17].

9. Eq. (38) can be used as well to compute inverse Fourier transform since  $\mathcal{F}^{-1}\{u(f)\} = \overline{\mathcal{F}\{u(f)\}}$  where  $\bar{a}$  is the complex conjugate of  $a$ .

odically. That is not at all the behavior of the Fourier transform of a function which is non-zero only in a finite interval of its domain. In that regard, eq. (38) is a much better approximation of the Fourier transform of a function for a much broader range of values of  $f = k\Delta f$ . At first glance, it may seem a minor point but sometimes (for example in a deconvolution process), it is important to have accurate values of the Fourier transform, even for values of  $f$  clearly outside the main characteristic bands of a function or a signal.

Another notable virtue of eq. (38), which is a by-product of the fact that it is a good approximation of the analytical Fourier transform, is that it gives, by setting  $k = 0$ , a very accurate numerical integration of the initial function between 0 and  $T$ . We will see that we can even obtain it between any limits of integration; but before let us turn our attention to differentiation.

An interesting attribute of the DFT is that it is exactly invertible. For example, applying the inverse operation (iDFT) to the sequence generated by  $F_{0,k}$  when  $k$  range from 0 to  $N - 1$  will regenerate exactly the sequence  $h_j^{(0)}$ ,  $j \in [0, N - 1] \subset \mathbb{N}$ . According to eq. (36),  $\mathbf{F}_b$  is completely determined and, when evaluated for all values of  $k$  from 0 to  $N - 1$ , it produces  $\theta$  different sequences of numbers which are actually very good approximations of the DFT of the derivatives of  $h_j^{(p)}$ , for all values of  $p \in [1, \theta]$ . Thus, applying the iDFT operation to each of these sequences will generate all the corresponding sequences  $h_j^{(p)}$ , which are the numerical derivatives of all orders from 1 to  $\theta$  of the initial function  $h_j^{(0)}$ . This means

---

10. The symmetry of those values of the FFT could not always be apparent, particularly when they are a superposition of values showing that symmetrical character. It may be the case if the initial function is a complex one, which consists in fact in two real functions.

that, provided  $\theta$  is adequately chosen, we can compute very accurately the derivatives of a digitized function or signal for all orders one could need.

It should be noted that high order derivatives are less accurate than low order ones because estimation of a derivative of order  $n$  is in fact calculated with an accuracy of order  $\theta - n + 1$ . Nevertheless, their contributions improve overall accuracy for any results based on them as, for example, interpolation, integration and Fourier transform. In the case where a high order derivative of a particular order  $n$  would be needed alone, with a high precision, one should use a value of  $\theta$  such that  $\theta - n + 1$  would be sufficient. Since, as we will see, errors decrease exponentially with  $\theta$ , it is unlikely that one could not meet any need. Having accurate derivatives of all needed orders means that high quality interpolation can be performed. It thus means too that we can accurately compute integrals between any limits of integration.

There is more; the derivatives calculated in that way exhibit a particularly interesting behavior. As it can be seen in eq. (9), the function  $h$  is approximated, for each interval  $[j\Delta t, (j+1)\Delta t]$ , by a different polynomial. For practical cases, the degree of these polynomials is limited to  $\theta$ . Even if it has not been explicitly stated as a requirement, the way equations are developed and linked together implies that, for each  $j \in [0, N-2]$  and for each  $p \in [0, \theta-1]$ , the  $p^{th}$  derivative, evaluated at  $(j+1)\Delta t$ , of the  $j^{th}$  approximating polynomial exactly equals the  $p^{th}$  derivative, evaluated at  $(j+1)\Delta t$ , of the  $(j+1)^{th}$  approximating polynomial. It means that at each point  $(j\Delta t, h(j\Delta t))$ , each approximating polynomial connect very smoothly to the next one. In other words, it means that if we consider all the  $N$  polynomials connected together as one function approximating the function  $h$ , then, this approximating function itself and all its derivatives for all orders up to  $\theta - 1$  are continuous.

Thus, we have here what is called a mathematical spline, with all its nice properties. (One more reason why  $\theta$  must be odd [17].)

Furthermore, eq. (32) is, as far as few conditions (discussed later) are considered, a very accurate method for specifying free parameters (often named “end conditions”) of such splines. Even if free parameters of splines can, to some extent, be arbitrary; it is admitted that the best end conditions are those based on the best available estimates of end derivatives [18]. Elements of matrix **B** in eq. (32) are closely related to those end derivatives, (see eq. (24)) and have the advantage that they are based on all points of the function  $h$  instead of only few points at each end; which is clearly in harmony with the concept of splines. In other words, when using eq. (36) (which contain **B** from eq. (32)) to compute polynomial splines, one doesn’t have to worry about end conditions, since they are automatically computed<sup>11</sup> and included in the process.

---

11. For example, if, between 0 and  $T$ ,  $h$  matches a periodic function, (with an integer number of period between 0 and  $T$ ) it would be convenient to use the so-called periodic splines. In this case, provided accuracy conditions are respected, eq. (32) will gives  $\mathbf{B} \cong 0$ , which is exactly the end conditions for periodic splines.



## Optimum parameters and precision

In order to cleverly use the present method, there remains two points that should be clarified: what are the best (or at least sufficient) values of  $N$  and  $\theta$  which should be used and what is the accuracy of the obtained results?

As it has been mentioned earlier, the link between  $\mathbf{F}_a$  and  $\mathbf{B}$  is not perfect. It is at its best for values of  $f = k\Delta f$  around  $N\Delta f/2$  and deteriorates gradually for values of  $f$  approaching  $\kappa N\Delta f$ ,  $\kappa \in \mathbb{R}$ . In eq. (31), increasing  $\theta$  tend to improve accuracy of  $\mathbf{B}$ , but, at the same time, it widens the interval  $\Omega$ . This implies the use of values of  $f$  deeper in the inaccurate region, which tends to lower accuracy. These two tendencies counteract and when they balance, we are at the optimum value of  $\theta$ , which will be denoted  $\theta_{opt}$ . Since  $\mathbf{B}$  is a matrix, it contains many terms. Thus, we must decide what is meant by the best value of  $\mathbf{B}$ . It could be the value that will give, at the end of the process, the most accurate Fourier transform, the better integral, or say the best third or fourth derivative. Here the user has a little latitude. Anyway, whatever criterion is chosen, it will not change very much the value of  $\theta_{opt}$ . So, firstly, we will choose a criterion that is closely related to the significance of the terms of  $\mathbf{B}$ , easy to understand, fast to compute and which is based on the predictive power of the Taylor polynomial formed with those terms of  $\mathbf{B}$ . The only extra cost is that we will have to sample<sup>12</sup> one more value ( $h_{N+1}$ ) from  $h$ . The main weakness of this criterion is that its accuracy is based on comparison with few points of  $h$ . For various reasons, for example

---

12. It is evident that this way of estimating errors would be inapplicable if it would be impossible to sample the extra point  $h_{N+1}$ . It is one reason why a much more general and powerfull, though more heavy on computing, method is subsequently presented.

if  $h$  is a noisy signal, one would prefer to use a more robust criterion. A general structure and example of such a robust criterion will be presented later.

The matrix  $\mathbf{B}$  from eq. (32) gives us accurate approximations of values of  $b_n$  for  $n \in [0, \theta - 1]$ . Let us recall eq. (24):

$$b_n = h^{(n)}(T) - h^{(n)}(0) = h_N^{(n)} - h_0^{(n)} \quad (39)$$

At this point, it is not possible, and not necessary, to separate the  $h_N^{(n)}$  from the  $h_0^{(n)}$ . According to the definition of Taylor expansion we can write:

$$h_{N+1} - h_1 = \sum_{p=0}^{\infty} \frac{(h_N^{(p)} - h_0^{(p)})(\Delta t)^p}{p!}, \quad p \in \mathbb{N} \quad (40)$$

Replacing  $(h_N^{(p)} - h_0^{(p)})$  in eq. (40) with the approximations given by  $\mathbf{B}$  and limiting the range of  $p$  to be between 0 and  $\theta - 1$ , will give us an approximation of  $(h_{N+1} - h_1)$ . The difference between this approximation of  $(h_{N+1} - h_1)$  and the exact value of  $(h_{N+1} - h_1)$  itself will be denoted and computed as follows:

$$e_{\theta} = \left| \left[ \left( \frac{(\Delta t)^0}{0!} \right) \left( \frac{(\Delta t)^1}{1!} \right) \dots \left( \frac{(\Delta t)^{(\theta-1)}}{(\theta-1)!} \right) \right] \mathbf{B} - (h_{N+1} - h_1) \right| \quad (41)$$

The way to use it is to begin with a low value of  $\theta$  and gradually increase it to obtain the desired accuracy. When  $e_{\theta}$  reaches its lowest value,  $\theta_{opt}$  is attained. At that point, if the required accuracy is not reached, the only way to go is to increase  $N$  and, if necessary, con-

tinue to increase  $\theta$  toward a new  $\theta_{opt}$ . The same kind of process could be used if one chooses to define  $\theta_{opt}$  in a different manner.

Let us now address the second point: the accuracy of the results. The value  $e_\theta$  is a good indication of the evolution of the accuracy of the results with the change of  $N$  and  $\theta$ . Generally, when  $e_\theta$  decreases, the average or maximum error of results (Fourier transform, integrals, derivatives...) decreases too and conversely. In that regard,  $e_\theta$  is significant and useful. However,  $e_\theta$  cannot be used as an approximation or as an upper bound for the errors. When, as it will be the case later, for purpose of comparison and testing, we use an initial function  $h$  for which the analytical form of result is known, the errors can be exactly determined. But when the analytical form of the initial function is unknown, as it is the case for experimental sampled data, or when analytical form of results is hardly or not obtainable, one still need approximations, or at least upper bounds, of the errors of computed numerical results. Let us handle it in a very simple way.

Since the present method is based on Taylor expansion, we will merely use these concepts for estimation of errors. Let us define  $R_\theta$  as any result obtained for a particular value of  $\theta$ ; and  $E_\theta$ , the error on  $R_\theta$ . Any result  $R_{\theta+2}$ , which involves two more terms than  $R_\theta$  in the Taylor expansions, can be considered, relatively to  $R_\theta$ , as being almost exact, provided  $\theta+2$  is not greater than  $\theta_{opt}$ . Thus, when  $\theta+2 \leq \theta_{opt}$ , an approximation of error can be given by:

$$E_\theta \cong O(R_{\theta+2} - R_\theta) \quad (42)$$

$O$  can be a function, a functional or an operator, depending on the nature of  $R_\theta$ , and on what is needed as  $E_\theta$ . Since  $R_{\theta+2}$  is necessary to estimate the error on  $R_\theta$ , it then appears better to use  $R_{\theta+2}$  instead of  $R_\theta$  as final result. Doing so will give, as long as  $\theta + 2 \leq \theta_{opt}$ , a more accurate final result. In this situation,  $E_\theta$  is very likely to highly overestimate error on  $R_{\theta+2}$ , so, it could be used, at that time, not as an approximation of error, but as an estimation of an upper bound of error on  $R_{\theta+2}$ . As we will see later, it works pretty well.

As mentioned earlier, sometimes one could prefer a more robust method than that using eq. (41) to estimate  $\theta_{opt}$ . Such a method involving particular instance of  $R_\theta$  and  $E_\theta$  can be built. An iterative process, similar to the one described for  $e_\theta$  can be used, but with  $E_\theta$  instead. If  $R_\theta$  is chosen to be a numerical function, appearing at the end of the computation chain, (Fourier transform, integral, n-derivative...) that use all the points of the initial function, and if  $O$  is based on average properties, one obtains a robust and safe method. However, it should be noted that the burden of computation will be higher than that entailing the sole use of eq. (41).

## Examples

In order to thoroughly test all aspects of the method, a non-exotic though particularly demanding function has been defined. It is:

$$f(t) = ( 2e^{-500(t-0.5)^2} + e^{-3.5t}, e^{-3t}\cos(6.25\pi t) ), \quad T = 1 \quad (43)$$

This a complex function. It is not band limited in the Fourier transform domain. The real part is a highly peaked Gaussian, slanted by an exponential to avoid symmetrical error cancellation. The imaginary part is a multiple cycle damped oscillation with a non-integer number of cycle between 0 and  $T$ . Both parts are non-zero and have non-zero  $n$ -derivatives nor at 0 neither at  $T$ , thus producing abrupt discontinuities at 0 and  $T$  for the function itself as well as for any of its derivatives.

First, let us evaluate  $e_\theta$  with eq. (41) for few values of  $N$  and  $\theta$ . Table 1 shows the results. As expected, accuracy improves when  $N$  increases. The same is true for the increase of  $\theta$ , as long as  $\theta \leq \theta_{opt}$ . Furthermore, when  $\theta \leq \theta_{opt}$ , error decreases exponentially with  $\theta$ . As an example, for  $N = 256$ , we have approximately  $e_\theta = 2.784e^{-3.148\theta}$ . It means that with as few as 256 points, for a function as much demanding as  $h$  (eq. (43)), one can obtain results with more than 50 exact significant figures.

Table 2 presents results based on eq. (42) to obtain  $E_\theta$ . Comparison between theoretical and estimated errors shows that when  $\theta \leq \theta_{opt}$ , the estimated errors are very good approximations of theoretical errors. And, as discussed earlier, if one chooses to use  $R_{\theta+2}$ , instead of  $R_\theta$ , as usable result, then  $E_\theta$  is here always a safe upper bound for error on  $R_{\theta+2}$ . In ad-

**Table 1:** Values of  $e_\theta$  (eq. (41)) for different values of  $N$  and  $\theta$ . Dark shaded cells indicate  $\theta_{opt}$ . Notation: 1.86 over (-09) stands for  $1.86 \times 10^{-9}$ . Computations were performed with 200 digits. Results are for function  $h$  as defined by eq. (43).

		$\theta$											
		1	3	5	7	9	11	13	15	31	43	49	51
$N$	32	3.89 (-01)	1.46 (02)										
	64	1.23 (-01)	1.15 (-02)	1.41 (-01)									
	128	5.39 (-02)	1.35 (-03)	2.03 (-05)	2.15 (-07)	1.86 (-09)	1.42 (-11)	9.61 (-14)	4.65 (-13)				
	256						6.71 (-15)	1.12 (-17)	1.63 (-20)	1.75 (-26)	2.76 (-58)	5.08 (-59)	1.38 (-56)

dition, table 2 exhibits results of an alternate way, based on eq. (42), to deduce  $\theta_{opt}$ , that gives almost the same value as the previous one, obtained with eq. (41).

**Table 2:** Values of  $E_\theta$  (Eq. (42)) with  $R_\theta$  being the numerical first derivative of  $h(t)$ , with  $N = 128$ .  $O$  here is a functional defined as the average on the 128 points of  $|R_{\theta+2} - R_\theta|$ . The theoretical error is computed in the same way but with the exact first derivative in place of  $R_{\theta+2}$ . Dark shaded cell indicates  $\theta_{opt}$ . Computations were performed with 50 digits. Results are for function  $h$  as defined by eq. (43).

		$\theta$								
		1	3	5	7	9	11	13	15	17
$E_\theta$ (Eq. (42))		8.7752 (-01)	6.88 (-04)	8.98 (-06)	1.53 (-07)	3.52 (-09)	9.67 (-11)	3.13 (-12)	2.06 (-13)	4.18 (-11)
Theoretical error.		8.7759 (-01)	6.95 (-04)	9.12 (-06)	1.56 (-07)	3.61 (-09)	9.97 (-11)	3.23 (-12)	1.35 (-13)	8.60 (-14)



Suppose one needs to calculate Fourier transform, integral and first few orders derivatives of  $h$ , with  $e_\theta$  around  $10^{-20}$ . Table 1 shows that  $N = 256$  and  $\theta = 15$  will fulfil these requirements. Table 3 shows the errors on the results for the Fourier transform for two different ranges of  $f = k\Delta f$ . As discussed earlier, we see that eq. (38) gives very good results in any range. Results for DFT are presented for purpose of comparison. To obtain, with DFT, even in a reduced range, a level of accuracy comparable to the one obtained with eq. (38), one should use  $N \cong 10^{20}$  instead of  $N = 256$  and it would need about  $10^9$  years of computation time instead of few seconds or less.

**Table 3:** Errors for different operations obtained with  $N = 256$  and  $\theta = 15$ . Computations were performed with 200 digits. Results are for function  $h$  as defined by eq. (43).

	Fourier transform (eq. (38))	Fourier transform (eq. (38))	DFT	Numerical integral
	$k \in [0, N - 1]$	$k \in [N, 2N]$	$k \in [0, N/2]$	$k \in [0, N]$
Average absolute error	3.20 (-23)	1.18 (-24)	2.82 (-3)	4.91 (-24)
Relative error	3.84 (-21)	2.00 (-21)	1.79 (-1)	1.71 (-23)

Since eq. (36) permits to obtain values of derivatives of all order up to  $\theta$  for each point, then interpolation by different Taylor polynomials between each two points can be done. It thus enables us to integrate between any limits. For simplicity, we restrict computation of integral between 0 and  $k\Delta t$ , for each value of  $k \in [0, N]$ . Results have been compared, to estimate errors, with the results from analytical integration between the same limits. Table 3 shows these results.

In table 3 and subsequently, average absolute errors are the averages, on the indicated numbers of points, of the absolute values of the differences between computed results (with the indicated method) and corresponding analytical ones. Relative errors are the average absolute errors divided by the average of the absolute value of the corresponding results.

Table 4 shows errors on numerical derivatives for all orders up to  $\theta$ . As expected, accuracy is very good even if it slowly decreases at higher orders. Note that because higher order derivatives take very large values, the average absolute errors may seem large but, relative errors, even for higher orders, always keep low values. For example, for fourteenth and thirteenth derivatives, relative errors are still below 2% and 0.023% respectively.

**Table 4:** Errors on numerical derivatives for all orders up to  $\theta$ . Note that even if  $\theta$  must be odd, derivatives are available and accurate for even order as well. Computations were performed with 200 digits.  $N = 256$ . Results are for function  $h$  as defined by eq. (43).

Order	1	2	3	4	5	6	7	8
Average absolute error	8.13 (-19)	1.31 (-15)	1.55 (-12)	3.00 (-9)	3.22 (-6)	6.90 (-3)	6.49 (0)	1.61 (4)
Relative error	1.05 (-19)	5.96 (-18)	1.74 (-16)	6.45 (-15)	1.07 (-13)	3.15 (-12)	3.69 (-11)	1.06 (-9)
Order	9	10	11	12	13	14	15	
Average absolute error	1.26 (7)	3.88 (10)	2.88 (13)	9.99 (16)	3.68 (19)	3.76 (23)	5.71 (26)	
Relative error	8.99 (-9)	2.86 (-7)	1.63 (-6)	6.64 (-5)	2.22 (-4)	1.94 (-2)	2.45 (-1)	

Let us now look briefly at connection spline conditions. For each  $j \in [0, N - 1]$ , let us form a Taylor polynomial with all  $h_j^{(p)}$ ,  $p \in [0, \theta]$ . Since these polynomials are polynomial splines, it is expected that for each  $j \in [0, N - 2]$  and each  $p \in [0, \theta - 1]$ , the  $p^{th}$  derivative,



evaluated at  $(j+1)\Delta t$ , of the  $j^{th}$  interpolating polynomial exactly equals the  $p^{th}$  derivative, evaluated at  $(j+1)\Delta t$ , of the  $(j+1)^{th}$  interpolating polynomial. The derivatives of all orders from 0 to  $\theta$ , for each  $j \in [0, N-1]$ , of these polynomials have been computed. Table 5 shows, among all junction points, the maximum discrepancy of the derivatives, at the connections points, of the joining polynomials; and that, for each orders from 0 to  $\theta$ . It is evident from these results that spline conditions are fulfilled.

**Table 5:** Connection spline condition. A.M.D. stand for absolute maximum discrepancy. Notation: 3 (-99) stand for  $3 \times 10^{-99}$ . Computations were performed with 100 digits.  $N = 256$ . Results are for function  $h$  as defined by eq. (43).

Order	0	1	2	3	4	5	6	7
A.M.D.	3 (-99)	6 (-99)	2 (-96)	2 (-94)	3 (-92)	5 (-89)	3 (-86)	3 (-83)
Order	8	9	10	11	12	13	14	15
A.M.D.	3 (-80)	5 (-77)	5 (-74)	4 (-71)	3 (-68)	2 (-65)	8 (-62)	3 (16)

Even if relatively good results have been obtained for eq. (43), they do not fairly reveal the full power of the method since that function has been purposely made difficult. Let us now present, without much discussion, the same kind of results as those shown in tables 3 and 4 but, this time, with the following slightly less demanding function:

$$h(t) = e^{-2t} \quad (44)$$

In spite of its apparent simplicity, it is not that trivial. It is still not band limited in the Fourier transform domain. The function and all its derivatives have abrupt discontinuities at both ends. It decreases rapidly and, as any exponential, it is always slanted on the same side. It is less demanding than eq. (43) in that it is a real function without peaks or fast oscillations.

Results shown in table 6 reveal that with as few as 256 points, absolute and relative errors on the Fourier transform are smaller than  $10^{-43}$ , for any range of  $k$ . For comparison, in its reduced range, the error on the DFT is 19%. The errors for numerical integral are lower than  $10^{-45}$ . Table 7 shows very good results for derivatives too.

**Table 6:** Errors for different operations obtained with  $N = 256$  and  $\theta = 15$ . Computations were performed with 200 digits. Results are for function  $h$  as defined by eq. (44).

	Fourier transform (eq. (38))	Fourier transform (eq. (38))	DFT	Numerical integral
	$k \in [0, N - 1]$	$k \in [N, 2N]$	$k \in [0, N/2]$	$k \in [0, N]$
Average absolute error	2.48 (-46)	1.03 (-49)	1.78 (-3)	1.78 (-46)
Relative error	5.02 (-44)	2.76 (-46)	1.90 (-1)	6.29 (-46)

Results for connection spline condition for eq. (44) are not presented here since they are similar though slightly better than those shown in table 5.

Results obtained with eq. (44) are indeed good for Fourier transform, integrals and for the first few order derivatives. Even for higher orders derivatives it is not so bad. However, all results presented until now are for functions sampled with not more than 256 points. For modern computers, digital storage oscilloscopes or data acquisition systems it is a very conservative number. Table 8 present results showing that Fourier transform, integrals and first few order derivatives can be obtained with much more than 100 exact significant figures.

**Table 7:** Errors on numerical derivatives for all orders up to  $\theta$ . Note that even if  $\theta$  must be odd, derivatives are available and accurate for even order as well. Computations were performed with 200 digits.  $N = 256$ . Results are for function  $h$  as defined by eq. (44).

Order	1	2	3	4	5	6	7	8
Average absolute error	3.70 (-42)	4.62 (-40)	2.39 (-36)	6.85 (-34)	1.55 (-30)	9.49 (-28)	1.01 (-24)	1.61 (-21)
Relative error	4.26 (-42)	2.60 (-40)	6.89 (-37)	9.86 (-35)	1.12 (-31)	3.42 (-29)	1.81 (-26)	1.45 (-23)
Order	9	10	11	12	13	14	15	
Average absolute error	6.61 (-19)	3.53 (-15)	4.46 (-13)	9.27 (-9)	3.21 (-7)	3.61 (-2)	55.4 (0)	
Relative error	2.98 (-21)	7.93 (-18)	5.02 (-16)	5.21 (-12)	9.04 (-11)	5.07 (-6)	3.90 (-3)	

**Table 8:** Value of  $e_\theta$  (eq. (41)). Computations were performed with 220 and 300 digits respectively. Results are for function  $h$  as defined by eq. (44).

$N = 4096$	$\theta = 49$	$e_\theta = 1.12 \times 10^{-103}$
$N = 16384$	$\theta = 55$	$e_\theta = 9.13 \times 10^{-140}$

## Computation time

Computation time may become a critical issue in many circumstances. Table 9 shows, for example, relative computation time for Fourier transform, once symbolic expression for eq. (38) is established. These results may vary depending on the way formulas are simplified, on how redundant evaluations are eliminated and on what kind of software and hardware are used<sup>13</sup>. Since eq. (38) must be evaluated for each value of  $k$ , computation time will depend linearly on  $N$ . Relative time shown in table 9 are in very good accordance with a quadratic expression depending on the order  $\theta$ . In this particular case, it is:  $t_{rel} = 0.55 + 0.41\theta^2$ . Accuracy increases exponentially while computation time increases quadratically, so, as long as one can afford it, the trade-off is beneficial. As an example, for  $\theta = 5$ , computation time is only ten times greater than that of the FFT while accuracy is  $10^8$  times better. For  $\theta = 15$  it takes 100 times more time to increase accuracy by a factor of  $10^{42}$ .

**Table 9:** Relative computation time for Fourier transform (eq. (38)) with  $N = 256$ . Unit time is that required to perform a FFT. Computations were performed with 25 digits. Results are for function  $h$  as defined by eq. (44).

		$\theta$								
		FFT	1	3	5	7	9	11	13	15
$t_{rel}$	1	1.3	3.6	9.6	22	33	46	69	92	

13. All symbolic and numerical computations have been done with the software MAPLE V, mainly with a desktop computer MACINTOSH LC3. For computations necessitating more memory a mainframe IBM RISC SGR2564 has been used. MAPLE V includes routines that convert formulas in C or Fortran and optimize them. Transferring computations in a fast compiled program is then an easy task.

## Comparisons

In this section we present comparisons between four different methods intended for computation of Fourier transform. The first method is described in [13]. It is based on correction applied to DFT. It gives exact result for a particular kind of functions that are said to be canonical-k. For others functions which are not canonical-k the method can still improve the results. For short we will refer to it as the “DFT+” method. The second method is described in [12]. It is based on local interpolation by Lagrange polynomials. Let us name it “LagPol” method. The third method is described in [7, 11]. This method “...use the efficiency of trapezoidal rule (using the FFT) in conjunction with the main features of the MIPS method.”[10]. MIPS stand for Möbius Inversion of the Poisson Summation formula. So, for brevity we will refer to it as the “MIPS” method. The fourth method in these comparisons is the method presented and developed in the present paper. We will designate it simply by “Present method”.

Table 10 shows results for functions  $e^t$  and  $e^{-2t}$  for  $t \in [0,1]$  and 0 elsewhere. Since these functions are of continuity order 0, the best results given by the DFT+ method should be, and indeed are those for  $\theta = 0$ . We see that for others values of  $\theta$ , the corrections do not really improve results, which is an expected behavior, given the nature of functions. LagPol method, for  $\theta = 0$ , give similar result as DFT+ but accuracy improve rapidly with the order as it is the case too with the Present method. Note that only even orders have been used with LagPol for reasons of symmetry.

Table 11 shows result for function  $h$  as defined in eq. (43). In this case, DFT+ method does not improve results due to the fact that  $h$  is too far from being a canonical function of any order. LagPol and Present method gives excellent results which are almost identical for low values of  $\theta$ . However, with Lagpol method, the improvement due to an increase of order



**Table 10:** Average absolute errors for different methods of Fourier transform applied on indicated functions. In the case of DFT+,  $\theta$  represents the order of correction for a canonical- $\theta$  function. For LagPol and Present method,  $\theta$  is the order of the system.  $N = 256$ ,  $T = 1$ . Computations were performed with 100 digits.

		$\theta$			
		0	1	2	3
DFT+	$e^t$	2.80 (-05)	2.19 (-03)	1.18 (-03)	1.63 (-03)
	$e^{-2t}$	2.01 (-05)	8.04 (-04)	4.36 (-04)	6.00 (-04)
LagPol	$e^t$	2.93 (-05)		2.05 (-10)	
	$e^{-2t}$	2.28 (-05)		1.82 (-10)	
Present method	$e^t$		4.72 (-06)		1.21 (-11)
	$e^{-2t}$		4.76 (-06)		4.86 (-11)
DFT	$e^t$	1.52 (-03)			
	$e^{-2t}$	5.59 (-04)			

begins to weaken at  $\theta \approx 5$ , while the accuracy continue to improve exponentially with the Present method. It depends on the fact that LagPol is a local method and that at high orders, as it is well known, Lagrange polynomials may not behave very well for good interpolation. Moreover a non-symmetrical scheme must be used for end conditions. On the other hand, Present method is, as discussed earlier, related to polynomial splines which are global in nature and particularly robust and well behaved for interpolation. Furthermore, with Present method, end conditions are globally and symmetrically determined.

**Table 11:** Average absolute errors for different methods of Fourier transform applied on function  $h$  as defined by eq. (43). In the case of DFT+,  $\theta$  represents the order of correction for a canonical- $\theta$  function. For LagPol and Present method,  $\theta$  is the order of the system.  $N = 256$ ,  $T = 1$ . Computations were performed with 100 digits.

	$\theta$											
	0	1	2	3	4	5	6	7	8	9	10	11
DFT+	1.16 (-02)	1.09 (-02)	1.12 (-02)	1.10 (-02)								
LagPol	3.06 (-04)		7.17 (-07)		7.51 (-09)		1.40 (-10)		3.64 (-12)		1.22 (-13)	
Present method		1.70 (-05)		2.55 (-08)		5.69 (-11)		1.17 (-13)		2.73 (-16)		9.07 (-19)
DFT:	1.10 (-02)											

Table 12 shows result for MIPS and Present methods applied on nine out of eleven functions listed in [10]; two of them, not continuous in  $[0,1]$  were not used. In [10], MIPS method has been used with 119 points on the average. To obtain, with Present method, similar or better results than those obtained with MIPS, a relatively high value of  $\theta$  must be used. With  $\theta = 17$  and  $N = 128 \approx 119$ , Present method is, for all but one function, many orders of magnitude more accurate than MIPS method. For  $N = 512$ , results are indeed still better, with, this time, no exceptions. However, one should absolutely note that limited accuracy in reported results for MIPS method (extracted from [10]) is not necessarily due to limitations of the MIPS method itself. It may be due to the limitations of the machine for which the smallest possible difference between the mantissa of two different numbers is about  $2 \times 10^{-18}$ . But, on the other hand, for function  $e^{7t}$ , which is a simple but demanding function, the MIPS method gives, for  $N = 109$ , an error of  $10^{-13}$ , which is well below the epsilon of the ma-

chine. By comparisons, for this function, with  $N = 64$ , Present method gives an error of  $10^{-21}$ .

**Table 12:** Average absolute errors for different methods of Fourier transform applied on indicated functions.  $T = 1$ ,  $p = 6$ . For Present method,  $\theta$  is the order of the system. Computations were performed with 100 digits except for results marked by \* for which 150 digits were used.

	MIPS	Present method $\theta = 17$	
		$N = 128$	$N = 512$
$e^t$	9.8 (-20)	3.22 (-44)	4.18 (-55)
$e^{4t}$	7.0 (-18)	1.72 (-32)	2.24 (-43)
$e^{7t}$	3.9 (-13)	4.75 (-27)	6.20 (-38)
$t - 1/2$	2.6 (-20)	1.14 (-84)	5.10 (-125)*
$t^2 - t + 1/6$	3.2 (-21)	3.70 (-86)	1.04 (-125)*
$\frac{1 - p^2}{1 - 2p \cos(2\pi t) + p^2}$	6.7 (-19)	9.45 (-29)	1.50 (-40)
$\ln(1 - 2p \cos(2\pi t) + p^2)$	3.1 (-18)	8.63 (-30)	1.36 (-41)
$\operatorname{atan}\left(\frac{2p \sin(2\pi t)}{1 - p^2}\right)$	2.8 (-19)	3.95 (-30)	3.84 (-43)
$e^{p \cos(2\pi t)} \cos(p \sin(2\pi t))$	9.2 (-20)	6.80 (-18)	2.57 (-30)
$e^{7t}$	3.9 (-13)	1.82 (-21) for $N = 64$	



## Conclusion

The present method can be conceived in few different manners. For example, it can be seen in the same logical way as it has been developed; a method to approximate the Fourier transform, but being able to give many other interesting results. Otherwise it could be thought as a method that use concepts related to Fourier transform to generate polynomial splines of any order with their end conditions, from which we can extract the other results. Whatever the way we choose to anchor it in our mind, it can be seen as an integrated and modular method at once. Eq. (38) can be used to establish, at an analytical level, the formulas to find numerical Fourier transform, without having to compute any derivatives or polynomial splines. On the other hand, eq. (36) can be used to obtain numerical derivatives of any order and polynomial splines of any order too, without having to compute Fourier transform. Eq. (32) is always involved in these cases. However, it can be used alone to give end conditions for any other numerical method that could need it. Furthermore, accurate numerical integrals can be obtained from Fourier transform or from integration of polynomial splines of required order.

All these results can be quickly obtained with much precision from a single and concise method developed without highly sophisticated mathematical concepts. The structure of equations permits to easily estimate errors or upper bounds on errors for any of the aforementioned results. Linked to the way errors are estimated, a general scheme to determine optimum parameters has been presented. It thus permits to automatically find the required parameters to meet a prescribed accuracy.

Finally, comparisons between many different methods have shown that if one need to compute accurately numerical Fourier transform, (and/or splines, n-derivatives, integrals...) the method presented in this paper should seriously be considered, not only for its power but for its simplicity, easiness to program and use and its versatility.

## **Acknowledgments**

My grateful acknowledgments go to Professor François Boucher who kindly revised all the manuscript. His invaluable support all along this work has been greatly appreciated.

## Bibliography

1. E. O. BRIGHAM. The Fast Fourier Transform. Prentice-Hall. Englewood Cliffs. 1974. pp.11, 98.
2. J. SCHUTTE. Rev. Sci. Instrum. **52(3)**, 400 (1981).
3. S. MAKINEN, Rev. Sci. Instrum. **53(5)**, 627 (1982).
4. S. SORELLA and S.K. GHOSH. Rev. Sci. Instrum. **55(8)**, 1348 (1984).
5. M. FROEYEN and L. HELLEMANS. Rev. Sci. Instrum. **56(12)**, 2325 (1985).
6. J. N. LYNESS. Math. Comp. **24**, 101 (1970).
7. J. N. LYNESS. Math. Comp. **25(113)**, 59 (1971).
8. J. N. LYNESS. Math. Comp. **28(125)**, 81 (1974).
9. J. N. LYNESS. J. Comput. Phys. **54(1)**, 57 (1984).
10. G. GIUNTA and A. MURLI. ACM Transactions on Mathematical Software, **13(1)**, 97 (1987).
11. J. N. LYNESS. QUG1: Trigonometric Fourier Coefficients. TM 370, Argonne National Laboratory. Argonne. 1981.
12. W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, B. P. FLANNERY. Numerical Recipes in FORTRAN. (2nd Ed.) Cambridge. New York. 1992. p.577.
13. N. MORRISON. Introduction to Fourier Analysis. Wiley-Interscience. New York. 1994. Chap. 13.
14. E. SCHIEFERDECKER and K. STREHLKE. *In* Fundamentals of Mathematics. Vol. 3. Analysis. H. BEHNKE, F. BACHMANN, K. FLADT, W. SÜSS, S.H. GOULD. The MIT Press. Cambridge. 1986. p. 80.
15. E. BUTKOV. Mathematical Physics. Addison-Wesley. Reading. 1968. p. 226.
16. E. KREYSZIG. Advanced Engineering Mathematics. (Fifth Ed.) Wiley. New York. 1983. p. A-55
17. J. H. AHLBERG, E. N. NILSON, J. L. WALSH. The Theory of Splines and Their Applications. Academic Press. New York. 1967. p. 109.

18. C. F. GERALD, P. O. WHEATLEY. Applied Numerical Analysis. (Fourth Ed.) Addison-Wesley. Reading. 1989. p. 211.

## **DEUXIEME ARTICLE**

**Fourier Transform deconvolution  
and  
time-resolved photoacoustic spectroscopy**

NORMAND BEAUDOIN and FRANÇOIS BOUCHER,  
*Groupe de Recherche sur l'Énergie et l'Information Biomoléculaire,  
Département de Chimie-Biologie,  
Université du Québec à Trois-Rivières,  
3351 boulevard des Forges, C.P. 500, Trois-Rivières,  
Québec, Canada. G9A 5H7  
(819) 376-5170*

Normand\_Beaudoin@UQTR.UQuebec.ca

Francois\_Boucher@UQTR.UQuebec.ca

02.90 Other topics in mathematical methods in physics.

06.50 Data handling and computation.

06.90 Other topics in measurement science, general laboratory  
techniques and instrumentation systems.

07.90 Other topics in specialized instrumentation.

87.80 Biophysical instrumentation and techniques.

August 5, 1999.

## **Abstract**

In this paper, we discuss the use of Fourier transform for deconvolution of real, digitized and noisy signals, in general, and more particularly, those issued from time-resolved photoacoustic spectroscopy (TRPAS) systems. Emphasis is put on the necessary conditions to obtain correct Fourier transform deconvolution results. Simulations with analytically known functions with and without noise and controlled test with real experimental signals show that a proper deconvolution method can produce correct result, in amplitude as well as in phase. Furthermore, tests reveal that the two time-related limitations in TRPAS experiments, the time resolution, linked to the diameter of the laser beam and, the maximum duration of measurements, limited by the reflections of signals, can be removed. Another problem with TRPAS systems, not related to but exposed by deconvolution and due to the amplitude dissymmetry between the sources of acoustic signals of the reference compound and of the active signal, is addressed. Experiments performed with bacteriorhodopsin contained in purple membranes have shown that it is possible, with Fourier transform deconvolution, to obtain significant and plausible results. The equations for temperature discrimination between thermal and conformational volume changes have been established and applied to the results of deconvolution to obtain conformational volume changes of bacteriorhodopsin.

## Résumé

Cet article discute de la pertinence d'utiliser les transformées de Fourier pour déconvoluer des signaux numériques réels avec bruit, en général, et en particulier, ceux provenant des systèmes de spectroscopie photoacoustique résolus en temps (TRPAS). Les conditions nécessaires pour obtenir des résultats corrects lors de déconvolutions par transformées de Fourier sont discutées. Des simulations faites à partir de fonctions analytiques connues, avec et sans bruits, et des tests à partir de signaux expérimentaux réels et, par ailleurs, déterminés, ont montré qu'avec une méthode de déconvolution appropriée, il est possible d'obtenir les bons résultats, tant en amplitude qu'en phase. De plus, des tests révèlent que peuvent être levées les deux limitations reliées au temps lors d'expériences de TRPAS, soient la plus fine résolution en temps dépendant du diamètre du faisceau laser, et la durée maximale de mesure, limitée par les réflexions. Un autre problème des systèmes de TRPAS, non-relié aux déconvolutions mais mis en évidence par elles, et dû à la dissymétrie entre les sources des signaux acoustiques des solutions actives et de référence, est aussi discuté. D'autres tests faits avec des membranes pourpres de bactériorhodopsine ont montré qu'il est possible, avec des déconvolutions par transformée de Fourier, d'obtenir des résultats visibles et plausibles. Les équations pour la discrimination en température, entre les changements de volume thermiques et conformationnels, furent établies et appliquées aux résultats des déconvolutions, pour obtenir les changements de volume conformationnels de la bactériorhodopsine.



## Introduction

The principles, methods and techniques of the so-called time-resolved photoacoustic spectroscopy (TRPAS) have been properly described by many authors [19-25]. TRPAS is considered as an excellent method for the study of light-induced radiationless energy relaxations processes, particularly those involving enthalpy changes and/or conformational changes resulting in volume changes.

Fig. 9 shows the diagram of an elementary TRPAS system. The functioning of such a system can be succinctly described as follows. A short laser pulse excites a sample, producing or initiating the reaction which may result in thermal and/or conformational volume changes. These energy relaxation processes generate an acoustic wave in the sample which propagates in all the components of the cell and is picked up by the piezoelectric element. The resulting electric signal can be amplified, sometimes filtered, digitized and stored in an oscilloscope before being forwarded to the computer for further treatment and analysis. Fig. 12 shows such typical signals.

The kind of information usually needed is the time profile of the thermal and/or conformational volume changes generated in the excited part of the sample. However, the recorded signal is not a direct function of these physical changes in the sample. It is the reaction of the whole system to such changes. It is the convolution of the derivative of the changes with the characteristic function of the system. Consequently the information must be extracted from raw signals by different means of different level of sophistication, depending on what is needed. Sometimes, only the amplitude of the signal is used. By comparing the amplitude of the signal produced by the photoactive sample under investigation to the one produced by an inert reference compound (one having the same thermoelastic properties and the same absorbance as the photoactive sample but which convert almost instantly the light energy into heat

without subsequent photochemical reactions), many authors have reported interesting results about enthalpy and/or conformational volume changes [26-31]. They even have been able to extract time-related information by considering the integration time of the system which is related to the laser beam's diameter.

However, to give the full meaning of the words "time-resolved" in a TRPAS experiment, one has to resort to the so-called deconvolution process. In TRPAS, one of the most often used method for deconvolution is the iterative reconvolution with least-squares criteria. It has been successfully used by many authors to obtain information on diverse substances as for example benzophenone, myoglobin, yellow protein, rhodopsin and bacteriorhodopsin [19, 21-25, 32, 33]. Briefly this method consists in defining a function with free parameters (often a sum of exponentials) assumed to be a good model for the volume changes that are the source of the signals. The free parameters are given initial values and the function is convoluted with the characteristic function of the system (the one produced by the reference compound). The result of this convolution is compared with the signal issued from the active sample. Iteratively, the free parameters are optimized to reduce, in the least-squares sense, the distance between the two compared functions. This method has the disadvantages of its advantages. It inherently incorporates a filtering process. Hence, one can always obtain a result as a best fit. But, finding the correct function that will reflect the behavior of the sample may be difficult, particularly for complex systems like bacteriorhodopsin which dissipates its energy in a multistep photocyclic process [30, 33, 45].

A deconvolution process that would directly give the result signal function from the reference and photoactive samples would be preferable. This process exists; it is the deconvolution by Fourier transform. However, its difficulties and the conditions under which it is applicable limited its use. Before addressing these issues, let us examine three limitations of the TRPAS systems.

The shape of the acoustic pulse generated in a liquid sample by a laser pulse has been studied by many authors. When the duration of the laser pulse is sufficiently short, the main parameter that determines the time width of the acoustic pulse is the diameter of the laser beam. The width of the acoustic pulse corresponds to the time needed by an acoustic wave to travel, in the sample, a distance equal to the diameter of the laser beam [20, 34, 35]. In a TRPAS system it is called the integration time and it is often considered as the best time resolution of the system. Many authors mention that a pinhole aperture can be used to reduce this diameter and thus set the resolution time [21, 22, 27, 28, 30-32]. In principle, with a proper deconvolution process, such a limitation could be eliminated.

Another limitation is the time needed by the acoustic wave reflected against the walls of the cell to come back to the sensor. The superposition of these reflections indeed results in a more complex signal. It is often considered preferable to use only the early part of the signals to conduct the analysis [21, 25]. Doing so gives a simpler signal but it may prevent the observation of late components in the signal of photoactive sample. Here too, a deconvolution process should be able to overcome this difficulty.

Another difficulty with TRPAS, not mentioned in the literature, is the dissymmetry between the source of the acoustic signals originating from the reference and from the photoactive samples. A good reference compound should have the same absorbance and thermoacoustic properties as the photoactive sample and convert almost instantly the absorbed energy into heat. Nevertheless, since the number of absorbed photons decrease exponentially from the entering side of the laser beam in the cell through the outgoing side, the amplitude of the local dilatation of the reference solution, which is the source of the acoustic wave, is not uniform all along the cylinder swept, in the sample, by the laser beam. This amplitude will decrease exponentially too, from the entering side of the laser beam in the cell through the outgoing side. On the other hand, if the number of photons in the laser pulse is

sufficient to saturate the photoactive sample, then, the source of the active component of the acoustic signal will be of the same amplitude all along the excited cylinder. In other words, if, immediately after the laser flash, the active sample is saturated all along the excited cylinder, then all volume elements of the active sample in that cylinder will undergo the same photocycle and produce volume changes of the same amplitude all along the cylinder. Hence, the active component of the signal produced with the photoactive sample, which its acoustic source will be spatially uniform, will be deconvoluted by a reference signal which its acoustic source will be spatially non-uniform. As we will see, this discrepancy in the spatial profile of the amplitudes of the sources of the acoustic signals, between the reference signal and the active component of the photoactive signal, may affect the result of the deconvolution.

In this paper we discuss the conditions for a successful Fourier transform deconvolution process with the help of analytical functions and artificial noise, and with real signals and noise. These examples will help to discuss the three aforementioned limitations. Then the whole process will be applied to signals obtained from purple membranes. A mathematical development applicable to Fourier transform deconvolution results, intended for the temperature discrimination between thermal and conformational volume changes in such samples will be presented.

## Fourier transform deconvolution

The aim of deconvolution process is to retrieve information blurred by the measuring system. If  $I(t)$  is the characteristic function (impulse response function) of the system and  $D(t)$  the response of that system to a causal input function  $E(t)$ , the relationship between these three functions is:

$$D(t) = \int_0^t I(\tau)E(t-\tau)d\tau = \int_0^t I(t-\tau)E(\tau)d\tau = I(t) \otimes E(t) = E(t) \otimes I(t) \quad (45)$$

It is the so-called convolution of  $I(t)$  by  $E(t)$ . It is a commutative operation. When  $I(t)$  and  $E(t)$  are known, obtaining  $D(t)$  is straightforward. But in many situations, the available signals or functions are  $I(t)$  and  $D(t)$ , and  $E(t)$  must be calculated. Any process by which  $E(t)$  is obtained from  $I(t)$  and  $D(t)$  is called a deconvolution. Direct extraction of  $E(t)$  from the integrals in eq. (45) is not so easy. Fortunately, once the functions are Fourier transformed, the operation becomes simple, eq. (45) can be written as:

$$\mathcal{F}\{D(t)\} = \mathcal{F}\{I(t) \otimes E(t)\} = \mathcal{F}\{I(t)\} \mathcal{F}\{E(t)\} \quad (46)$$

A convolution in the time ( $t$ ) domain become a simple multiplication in the frequency ( $f$ ) domain. Hence, the deconvolution can be performed in the frequency domain as follow:

$$E(t) = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{D(t)\}}{\mathcal{F}\{I(t)\}} \right\} \quad (47)$$

At first glance, using eq. (47) for deconvolution seems straightforward and easy. Nevertheless, major difficulties appear when  $D(t)$  and  $I(t)$  are digitized and noisy signals.

As it is evident in eq. (47), deconvolution by Fourier transform involves a division in the frequency domain. The denominator function is the Fourier transform of the characteristic function of the system. If this denominator function departs from its correct value, it is evident that the deconvoluted signal will be greatly distorted. The same is true if the numerator function is not accurate, but the effect on the final result is less dramatic than for imprecisions in the denominator function, particularly for regions where this last function is close to zero. Departures from correct values in these functions have two sources; discrete Fourier transform (DFT) and noise. Let us discuss these two points separately.

Usually the Fourier transform of a function that has been digitized equidistantly is approximated by the so-called discrete Fourier transform (DFT). (These DFT are almost always computed via an algorithm of fast Fourier transform (FFT)). However, as it has been discussed by several authors [36-40], DFT is, in many situations, far from being adequate to act in place of Fourier transform. It is particularly the case for deconvolutions. Let us illustrate this with an example.

Defining the characteristic function of the system as,

$$I(t) = \begin{cases} e^{at} [b_1 \cos(2\pi f_1 t) + b_2 \sin(2\pi f_2 t)] & \text{for } t \in [0, T] \\ 0 & \text{elsewhere} \end{cases} \quad (48)$$

and the input function as,

$$E(t) = \begin{cases} cte^{dt} & \text{for } t \in [0, T] \\ 0 & \text{elsewhere} \end{cases} \quad (49)$$

where  $a$ ,  $b_i$ ,  $f_i$ ,  $c$  and  $d$  are constants. One can compute the convolution of eq. (48) and eq.

(49) to obtain the following somewhat elaborated expression for  $D(t)$ :

(50)

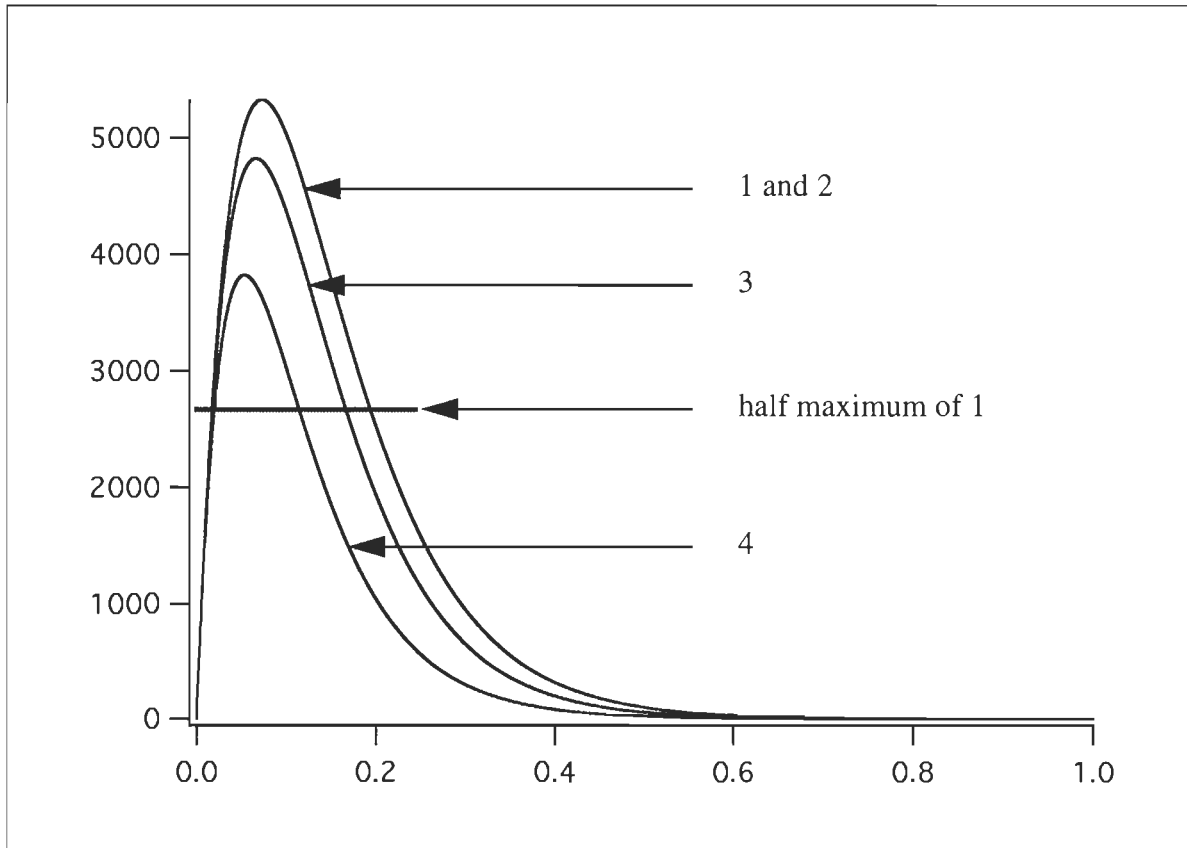
$$\begin{aligned}
 \omega_1 &= 2\pi f_1, \quad \omega_2 = 2\pi f_2, \quad \tau = a - d \\
 A_1 &= \tan\left(\frac{1}{2}\omega_1 t\right), \quad A_2 = \tan\left(\frac{1}{2}\omega_2 t\right) \\
 B_1 &= \omega_1^2 + \tau^2, \quad C_1 = \omega_1^2 - \tau^2 \\
 B_2 &= \omega_2^2 + \tau^2, \quad C_2 = \omega_2^2 - \tau^2 \\
 D_1(t) &= \frac{(C_1 A_1^2 + 4\omega_1 \tau A_1 - C_1)b_1 c e^{at}}{(1 + A_1^2)B_1^2} + \frac{(C_1 - \tau B_1 t)b_1 c e^{dt}}{B_1^2} \\
 D_2(t) &= \frac{2(\tau \omega_2 A_2^2 - C_2 A_2 - \tau \omega_2)b_2 c e^{at}}{(1 + A_2^2)B_2^2} + \frac{\omega_2(B_2 t + 2\tau)b_2 c e^{dt}}{B_2^2} \\
 D(t) &= D_1(t) + D_2(t)
 \end{aligned}$$

In the present example, the following numerical values for parameters will be used:  
 $a = \ln(10^{-3})$ ,  $b_1 = 1$ ,  $b_2 = 0.25$ ,  $f_1 = 100$ ,  $f_2 = 2000$ ,  $c = 2 \times 10^5$ ,  $d = 2a$  and  
 $T = 1$ . These parameters have been chosen such that the result of deconvolution may seem qualitatively correct, at first glance, if no comparison with a correct result is possible. It would indeed be possible to fix parameters to obtain better or worst deconvolution with DFT. The functions  $I(t)$  and  $D(t)$  are equidistantly digitized for a number of points  $N = 16384$ . Then the deconvolution is performed according with eq. (47), but with numerical approximations for Fourier transform and inverse Fourier transform instead of their exact values. An approximation of  $E(t)$  is thus obtained. Curve 4 of Fig. 7 shows the result of deconvolution with an ordinary DFT with a 4x 0-padding<sup>14</sup>. In fact, with the 0-padding, the DFT is applied

14. nx 0-padding means that many 0 are added at the end of a numerical function of  $N$  points to make it having  $nN$  points.

on  $4N = 65536$  points, as if the function were prolonged beyond  $[0, T]$ , up to  $[0, 4T]$ . So, the frequency resolution is improved, but not the time resolution. If we digitize the functions with four times more points ( $N = 65536$ ) in  $[0, T]$  to increase the time resolution, and still use a 4x 0-padding, (which means that the DFT is applied, this time, on  $4N = 261144$  points, in  $[0, 4T]$ ) then curve 3 of Fig. 7 is obtained. Compared to exact result for  $E(t)$  given by eq. (49) and represented by curve 1 of Fig. 7, these results are not very good, even if a huge number of points was used.

**Fig. 7.** Deconvolution of  $D(t)$  (eq. (50)) by  $I(t)$  (eq. (48)). Curve 1 shows the exact result,  $E(t)$  (eq. (49)). Curve 2, indistinguishable from curve 1, is the result obtained with the author's method. Curve 3 and 4 are obtained with DFT using different numbers of points.





To overcome this difficulty, one has to resort to a method that will give a better approximation to Fourier transform than does DFT. Any sufficiently accurate method can be used. Nevertheless, one should be very careful since, in virtue of the sensitivity due to the division in eq. (47), the chosen method should give a high (and preferably adjustable) level of accuracy. One of us developed such a method [40]. (We will refer to it as A.M. for author's method.) It involves a parameter, the order of the system, denoted  $\theta$ , which determines the accuracy of many numerical results including Fourier transform. Briefly, one way this method can be seen is as an implementation, in the Fourier transform domain, of spline interpolation of order  $\theta$  with intrinsic optimum end conditions. It appears to be particularly powerful in spite of its simplicity.

In our previous example, approximations of Fourier transform and inverse Fourier transform with that method (for  $\theta = 5$ ) give the result shown by curve 2 in Fig. 7. This curve is visually indistinguishable of curve 1, the exact result. Hence, to put numbers on comparisons, the errors on different curves are summarized in table 13.

Fig. 7 and table 13 clearly show the advantages of using numerical methods that approximate as much as possible the behavior of exact Fourier transform. To achieve good result for deconvolution with Fourier transform, it is almost mandatory to use such methods.

Let us now discuss the second aforementioned difficulty encountered with deconvolution by Fourier transform: noise. A division of a function, noisy or not, by another noisy function gives a noisy result. When the denominator function gets close to zero, there is almost an "explosion" of noise in the result of the division. This is exactly what is susceptible to happen with eq. (47). Very often a Fourier transform function exhibits a tail approaching zero for high frequency. Fortunately, a method giving very good results has been suggested to solve this problem [41]. Essentially it consists in replacing the high frequency parts of the noisy

**Table 13.** Errors in deconvolution with ordinary DFT and accurate algorithm of Fourier transform. In all cases a 4x 0-padding is used. A.M. stands for author's method. Relative RMS errors are computed relatively to exact result (eq. (49), curve 1 in Fig. 7). Relative error on peak's position is the departure of the considered peak from peak's position of the exact curve, relatively to full width at half maximum (fwhm) of curve 1.

	DFT $N = 16384$	DFT $N = 65536$	A. M. ( $\theta = 5$ ) $N = 16384$
Relative RMS error for $t \in [0, T]$	41.4 %	15.1 %	0.06 %
Relative RMS error for $t \mid E(t) > \frac{1}{2} \max(E(t))$	51.3 %	17.7 %	0.02 %
Relative error on peak's position	10.9 %	3.7 %	0.00 %

Fourier transform of initial functions by the Fourier transform of an exponential. It is indeed particularly relevant when the initial function is an exponential, and it may even work for more complex functions. Nevertheless, it may happen that the denominator function,  $\mathcal{F}\{I(t)\}$ , gets close to zero for low frequency regions, which is the case, for example, for an underdamped middle frequency oscillating signal. It may even happen that  $\mathcal{F}\{I(t)\}$  is almost zero for a relatively wide region between two other regions of significant non-zero values, which is the case for composite signals as for example the one defined by eq. (48). In these situations, if one does not want or cannot impose a functional form to the signals and still takes the advantages of direct deconvolution by Fourier transform, another approach must be attempted which should accommodate with that noisy division.

It is often said, and sometimes written [42], that the division in the frequency domain generates noise. One should prefer the more correct formulation of the authors of ref. [41]: "... the presence of this noise may give rise to a mathematically exact result that has no phys-

ical significance.” Outside the fact that there exists perfectly deterministic computer programs which generate pseudo-random numbers, it cannot be said that the division in Fourier transform deconvolution really “generates” noise. If one deeply adhere to that point of view, he will try to find an alternate way that would avoid the division and which should give the same result of the Fourier transform deconvolution but without the noise. This is not possible since a deconvolution algorithm cannot “know” what is the noise and what is the signal. In fact the “amplification” of the noise due to the division is not really an amplification. It should be understood that, according to the definition of convolution, for a particular  $t_0$ , the value  $E(t_0)$  affect  $D(t)$  for all  $t > t_0$ . Then, what one may call the noise on  $E(t)$  is due to the wild fluctuations that must experience  $E(t)$  to reproduce in details  $D(t)$  with its noise, when it drives a system whose characteristic function  $I(t)$  is itself noisy. Then what may appear as an undesirable behavior which seems to generate noise is in fact is the correct behavior of a good deconvolution system. Let us look at a possible solution of that noise problem.

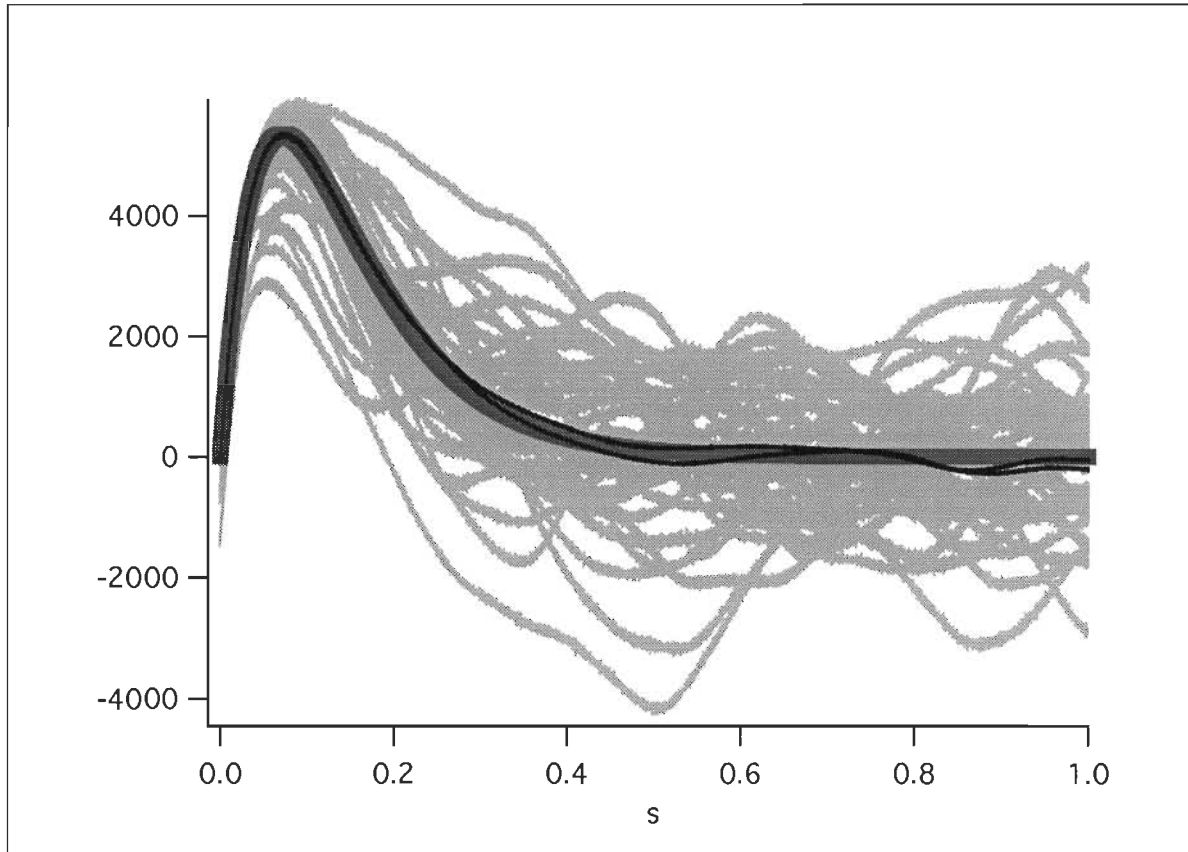
Suppose that a system has been designed to be as little noisy as possible and that all proper analog or digital techniques have been applied to each single signal, up to a point beyond which further filtering would begin to distort the needed information. In such a situation, if the results of Fourier transform deconvolution of several signals (all of them produced by a system always identically prepared) are not sufficiently similar, then the noise must be reduced in a manner that does not alter the information contained in the signals. Averaging several functions is probably one of the most simple and efficient method to do that. (The use of a robust estimator, like median, could be considered too.) An example will illustrate the efficiency of such a noise reduction process.

Let us use 64 functions  $I(t)$  and 64 functions  $D(t)$  as previously defined by eq. (48) and (50) respectively but, this time, with Gaussian noise added to each of them. The standard

deviation ( $\sigma$ ) of the noise is set to 0.0033 so that the peak amplitude ( $3\sigma$ ) of the noise is approximately 1% of the maximum value of the functions. On the average, on all the interval  $t \in [0, T]$  it gives a noise of 2.8% on  $I(t)$  and 4.7% on  $D(t)$ . One can perform the 64 deconvolutions and take the average of the results, or, take the average of the 64 functions  $I(t)$  and  $D(t)$  respectively and perform one deconvolution with these averages. The results are shown in Fig. 8. It clearly shows that individual deconvolutions (light shaded curves) can widely depart from the exact result. The average of deconvolutions (AD) and the deconvolution of the averages (DA) are almost identical and in good agreement with the exact (noiseless) result. It should be noted that the AD scheme appears more safe than the DA one. In the DA scheme, only one deconvolution is performed and, even if the noise of the initial functions is greatly reduced by the averaging, it may happen that the denominator function contains few points very close to zero that would spoil the result. In the AD scheme, 64 (in this example) deconvolutions are performed. It appears that 7 of them lead to result functions containing at least one point (among 16384) outside of 3 standard deviation from the average. These 7 functions have been removed and a new average computed. This AD scheme also allows to estimate the dispersion of results. Such a discrimination of too bad results due to noise in the denominator function, or estimation of dispersion are not possible in the DA scheme. (Even if it has not been necessary in this example, a variant of the AD scheme based on robust estimator like median instead of average could be used. Robust estimators are very efficient to almost completely eliminate the effect of very bad results [43].)

In this example, 64 functions  $I(t)$  and 64 functions  $D(t)$  have been used. With real experimental signals, generally,  $I(t)$  is obtained from a reference compound, and  $D(t)$  comes from the photoactive sample. Hence obtaining many  $I(t)$  may be much more easy than obtaining many  $D(t)$ , particularly when  $D(t)$  is produced by a fragile substance like, for example, rhodopsin, a protein that bleaches when illuminated. In such a situation, since it

**Fig. 8.** Deconvolution of noisy signals. The light shaded background is the stacking of the 64 deconvolutions of the noisy signals. The two, almost identical thin black curves are, for one the average of 57 of the 64 individual deconvolutions (AD), and the other is the deconvolution of the averages (DA) of the 64 initial noisy functions  $I(t)$  and  $D(t)$  respectively. The dark shaded thick curve is the theoretical noiseless result given by eq. (49).

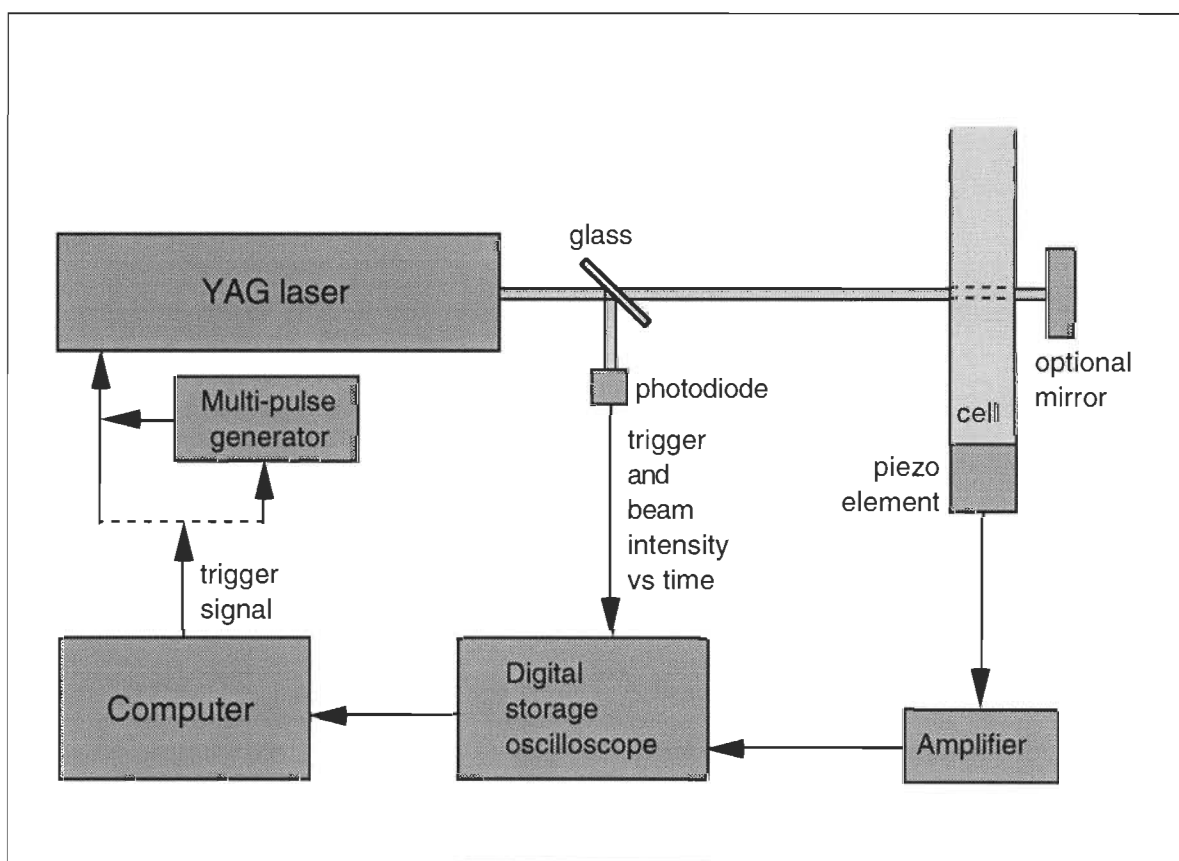


is the denominator function ( $\mathcal{F}\{I(t)\}$ ) which is mainly responsible for the noise in the deconvolution, one could consider to average the deconvolutions of the few available  $D(t)$  with many  $I(t)$  and still obtain a good result.

## Test experiment

Up to this point, only synthetic signals and artificial noise have been used. Let us now test the Fourier transform deconvolution process on real experimental signals produced by the simple TRPAS system depicted in Fig. 9 and briefly described hereafter.

**Fig. 9.** Schematics of the time-resolved photoacoustic spectroscopy system.



The light source is the second harmonic (532 nm) of a YAG laser (Lumonics JK Laser System 2000). From calorimetric measurements with aqueous dye, ( $A_{532} = 0.3$ ), we estimated that  $30 \pm 10$  mJ was deposited in the cell a each single laser flash, in the absence of

the back mirror. The multi-pulse generator was a home-made circuit that could fire the q-switch up to 3 times during one discharge of the pumping flash lamp of the laser. The laser beam excited the sample contained in a  $0.01 \times 0.01 \times 1$  m square glass tube. Its height was deliberately oversized to avoid the effect of irregularities on the surface of the liquid. In fact, multiple reflections of the acoustic wave on the walls of the cell proven to be reproducible, contrary to those on the air/liquid interface, particularly from one refill to another. With such a lengthy cell, reflections on the top of the sample were sufficiently delayed in time so that they reach the detector after the end of signal sampling. The disadvantage of this long cell is that the adjustment of temperature and its stability has always been problematic, though workable, at least for demonstration purpose. Signals were measured with a piezo-element ( $\text{PVF}_2$  piezoelectric plastic) wedged with a solid lead cylinder against the bottom of the cell, a quartz disc fixed with epoxy glue. They were fed into the amplifier (SR560 Stanford Research; high cutoff 1 MHz) and stored in the oscilloscope (Nicolet 4094C with 4180 plug-in) before being transferred to a desktop computer.

The system was operated in the following way. The multi-pulse generator was disabled and a single laser pulse (duration of 15 ns) sent in the sample. This was considered as a Dirac delta of power in the sample that produced the signal  $I(t)$ ; the characteristic function of the system<sup>15</sup>. Thirty two<sup>16</sup> of these signals were averaged to produce the final  $I(t)$ . A typical example of such a function is shown in Fig. 12. Then, the multi-pulse generator was turned on. It enabled the laser to fire 3 successive pulses of different amplitude. These 3 pulses constitute a function of power versus time that is in fact  $E(t)$  which in turn produced  $D(t)$ . Again 32 of these signals ( $D(t)$ ) were averaged to give the final  $D(t)$ . Due to the jitter of the generator, the second and third pulses of  $E(t)$  had an effective width of approximately 180 ns

---

15. We continue to use the names  $I(t)$ ,  $D(t)$  and  $E(t)$  in this context. It is clear that here, these functions are not the same as those defined by eq. (48), (49) and (50).

16. For all the experiments, the repetition rate of the laser pulses was one per two seconds (0,5 Hz).

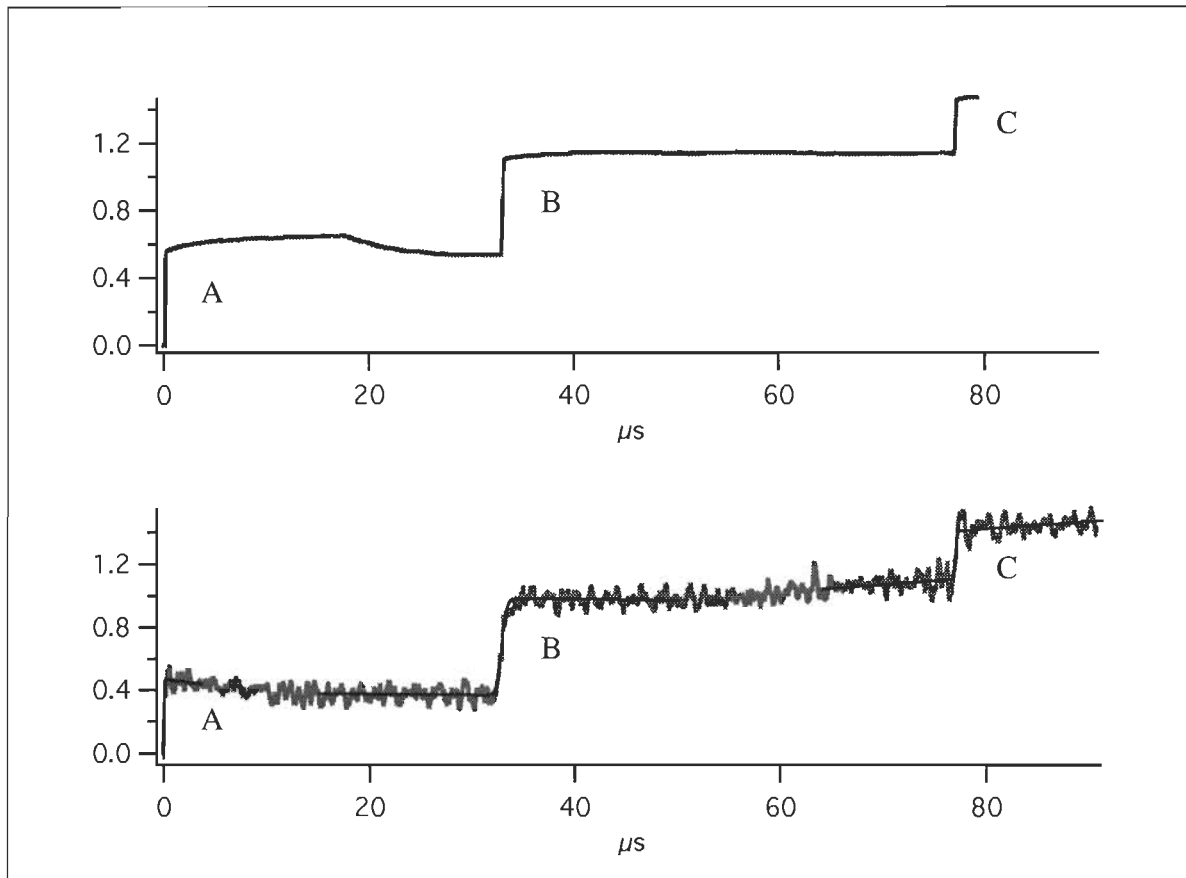
(instead of 15 ns) when averaged. The initially unknown  $E(t)$  will be recovered through the deconvolution of  $D(t)$  by  $I(t)$ . For purpose of testing and comparisons, the instantaneous power of the laser beam was probed via a fast photodiode. The signal from this photodiode gives  $aE(t)$ ,  $a$  being a constant. For comparisons we will consider the integral of the signals of the photodiode and the integral of the signals from deconvolutions. Doing so makes more clear and straightforward the comparisons since steps are more easy to compare than narrow peaks. Furthermore, the integral of power is energy; then the integral of  $E(t)$  shows the time profile of the amount of energy deposited into the sample. In Fig. 10, specimens of these profiles are shown. The upper one is from the photodiode. The lower one is from the deconvolution. Each steps are identified by a letter to facilitate the discussion.

Two identical (except for noise and jitter) signals have been recorded with the photodiode and 4 experiments have been conducted to do deconvolutions. Assuming that the laser pulses are Gaussian, an error function of the form  $a_0 + a_1 t + a_2 \text{erf}[(t - a_3)/a_4]$  has been fitted to each steps to precisely estimate their amplitudes, positions in time and durations. The amplitude of a step is  $2a_2$ , its position in time is  $a_3$ . The duration of each steps has been taken as the fwhm of the corresponding Gaussian pulse. It is  $1.67a_4$ . Table 14 presents the compilation of results drawn from the 6 aforementioned curves consisting in 6 steps from photodiode signals and 12 from deconvolution signals.

There is a very good match between the photodiode and deconvolution signals with respect to amplitude ratios and time position, as well. It constitute an a posteriori proof of the linearity of the system in amplitude and in phase and means that with a proper method, direct deconvolution of experimental signals is possible and can gives coherent results.



**Fig. 10.** Typical signals of  $\int aE(t)$  measured from the photodiode (upper graph) and  $\int E(t)$  obtained by deconvolution (lower graph). Parameter  $a$  has been adjusted to make the amplitudes comparable. In the lower part, the thin black curves show the results of the curve fitting of the error function to the steps.



Furthermore, with a cell width of 1 cm and a distance of 2 cm between the excited region of the sample and the sensor, there is a maximum delay of only 13  $\mu s$  between the laser pulse and the first reflections of the acoustic waves against the cell's walls onto the sensor. Table 14 shows that steps at 33  $\mu s$  and 77  $\mu s$  are very well resolved in time and amplitude. It means that the limitations, due to multiple reflections of the acoustic wave in the cell can be removed.

In the present experiment, with a laser beam diameter of 4.5 mm, the aforesaid limitation due to the diameter of the laser beam would have resulted in a best time resolution of 3

**Table 14.** Comparisons between direct photodiode measurement of the exciting function  $E(t)$  and its recovery by deconvolution. The letters A, B, C refer to corresponding labels in Fig. 10.

		B/A	C/A	C/B
Ratio of amplitudes	Photodiode	$1.03 \pm 0.05$	$0.61 \pm 0.02$	$0.59 \pm 0.03$
	Deconvolution	$1.05 \pm 0.10$	$0.55 \pm 0.06$	$0.53 \pm 0.06$
		A	B	C
Position in time [ $\mu s$ ]	Photodiode	0 (ref.)	$32.84 \pm 0.04$	$77.00 \pm 0.04$
	Deconvolution	$0.0000 \pm 0.0003$	$32.90 \pm 0.04$	$76.98 \pm 0.05$
Average of durations of pulses		$t_{fwhm} = 500 \pm 440$ ns		

$\mu s$ . The actual time resolution, as shown in table 14, is 6 times smaller. It means that this limitation too can be removed. So there is no need for small pinhole aperture any more. Hence it allows a larger sample volume to be excited. It is another virtue of deconvolution. Ideally there are no limitations in improving time resolution with deconvolution. Even the duration of the laser pulse could be accounted for in the deconvolution. In practice it depends on the possibility to obtain sufficiently noiseless signals.

The third aforementioned problem is due to the non-uniformity of the amplitude of the source of the acoustic waves along the cylinder swept, in the cell, by the laser beam. Let us look at the effect of the discrepancies in amplitude profiles of two real signals which, for purpose, have been made such that each one has its own non-uniformity. To produce two signals with different amplitude profiles with respect to the source of the acoustic wave along the illuminated cylinder, a mirror was put behind the sample as shown in Fig. 9. The first signal is produced with the reference compound ( $A_{532} = 0.3$ ) without the mirror. It is considered as  $I(t)$ . The second is produced with the same reference compound but with the mirror<sup>17</sup>. It is

17. These signals  $\{I(t) \text{ and } D(t)\}$  have been obtained with the laser's multi-pulse generator disabled.

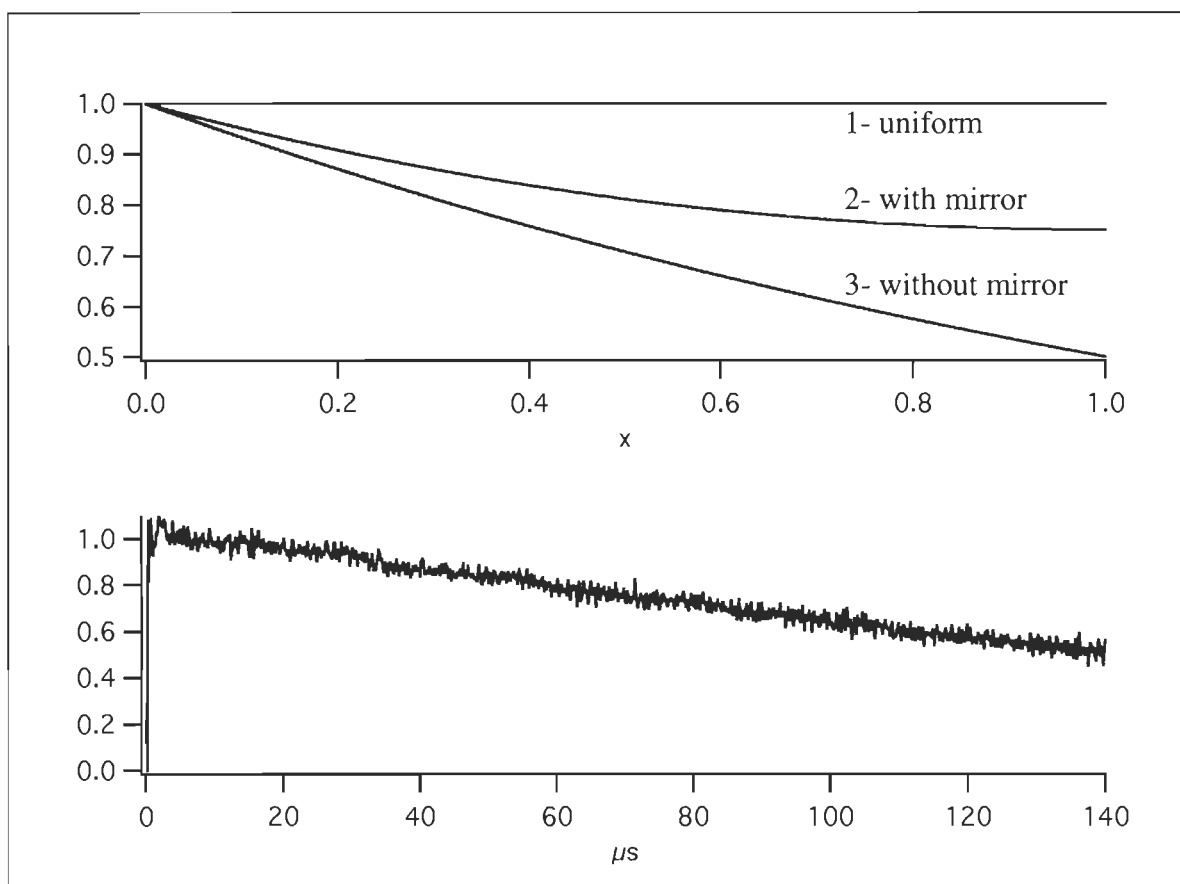
considered as  $D(t)$ . The amplitudes ( $A_I(x)$  and  $A_D(x)$ ) of the source of the signals that produce  $I(t)$  and  $D(t)$  are described respectively as follows:

$$\begin{aligned} A_I(x) &\propto 10^{-0.3x} \\ A_D(x) &\propto 10^{-0.3x} + 10^{-0.3(-x+2)} \end{aligned} \quad (51)$$

in which  $x \in [0, 1]$  cm is the depth in the cell taken from the entering side of the laser beam. The shapes of these amplitude profile are shown in the upper graph of Fig. 11. Since the important point here is the discrepancies between the profiles of the curves and not their absolute amplitude, they have been vertically translated to begin at the point  $(0, 1)$ , for comparisons purpose.

With the particular setting of the TRPAS system of Fig. 9, the integral of the deconvolution of the signal produced by the amplitude profile of curve 2 in Fig. 11 by the one produced by curve 3 gives the curve of the lower graph of Fig. 11. The result of the integral of the deconvolution of two identical functions is a Heaviside unit step function. Instead of that expected result, we obtain a step at  $t = 0$  followed by a negative slope line. It should be noted that even if such an effect has been made evident by the deconvolution applied to the whole signals instead of only to the first parts, it is not due to the deconvolution process itself. To minimize this effect, the mirror was always used for the reference and for the photoactive samples. In fact, when a saturating laser pulse is used, a photoequilibrium is established in a photoactive sample, resulting in an uniform source of the acoustic wave with respect to the conformational component of the signal. The effect, shown in the lower graph of Fig. 11, would be increased if a signal produced by a uniform source of acoustic wave, as curve 1 in Fig. 11 was deconvoluted by the one produced by curve 3, since the difference of profile between curve 1 and 3 is indeed greater than the one between curve 2 and 3. However, even if

**Fig. 11.** Upper graph: Amplitude profiles of the source of acoustic wave through the cell. Lower graph: Deconvolution result for two signals with non-uniform spatial distribution of the amplitude of their sources of acoustic wave.



the difference in profiles between curves 1 and 2 is less than the one between curves 1 and 3, it is nevertheless greater than between curves 2 and 3. So it is expected that conformational component of photoactive samples show a slope even more negative than that of the lower graph of Fig. 11.

## Experiment with purple membranes

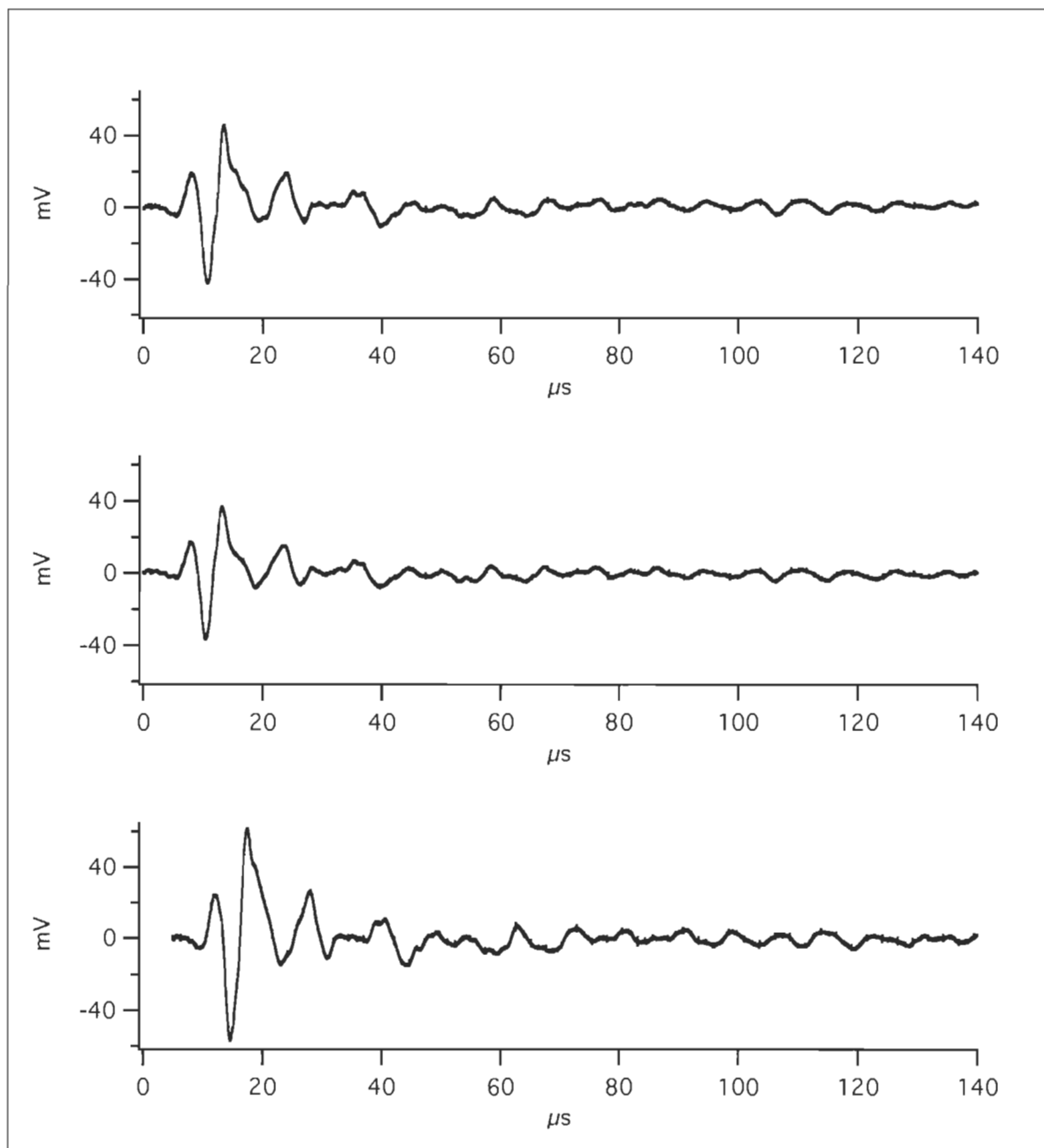
So far we have used computer generated functions and noise, or real but purposely tailored signals as those produced with the multi-pulse generator or with the “with and without mirror” configuration, applied to a reference compound. In order to check the reliability of the experimental set up and of the deconvolution process, a photoactive sample like bacteriorhodopsin is probably best suited as it has already been characterized by comparable methods. Indeed, thermal and structural volume changes which accompany the photocyclic activity of bacteriorhodopsin have already been analysed by time-resolved photoacoustic or photothermal methods [30-33, 44]. Considering the TRPAS systems shown in Fig. 9, gathering of accurate and stable data proved to be tedious. For example, the task of adjusting and maintaining the temperature of the cell was accomplished through stabilization of the temperature of the whole room. Nevertheless, it has been possible to obtain reproducible data from reference and photoactive samples at two temperatures.

Fig. 12 shows such typical raw signals at 11.5°C and 22.5°C. The first graph comes from the reference compound (aqueous black ink,  $A_{532}=0.3$ ). The second graph comes from the photoactive sample, purple membranes in pure water, with the same absorbance<sup>18</sup> at 11.5°C. The third graph comes from the photoactive sample too but, this time, at 22.5°C. At first glance, these three signals seem almost similar. Nevertheless a closer look reveals that they are everywhere slightly different. The most visible dissimilarity is in the amplitude. First, the photoactive sample at 11.5°C signal is less than that of the reference compound at the same temperature. It reveals the fact that a part of the light energy is stored by bacteriorhodopsin and is dissipated later instead of being immediately converted into heat. Second, the photoactive signal at 22.5°C is larger than the one at 11.5°C and even larger than the ref-

---

18. The quality and functionality of the samples of purple membranes have been established by many spectrophotometer scans and by many flash photolysis experiments.

**Fig. 12.** First graph: signal from the reference compound at 11.5°C. Second graph: signal from active sample at 11.5°C. Third graph: signal from active sample at 22.5°. Each signal is the average of 32 raw signals produced by 32 laser pulses. All graphs are at the same scale.



erence signal at 11.5°C. It is due to the fact that the thermal expansion coefficient of water is larger at 22.5°C than at 11.5°C. Nevertheless, to completely reveal the underlying causes of

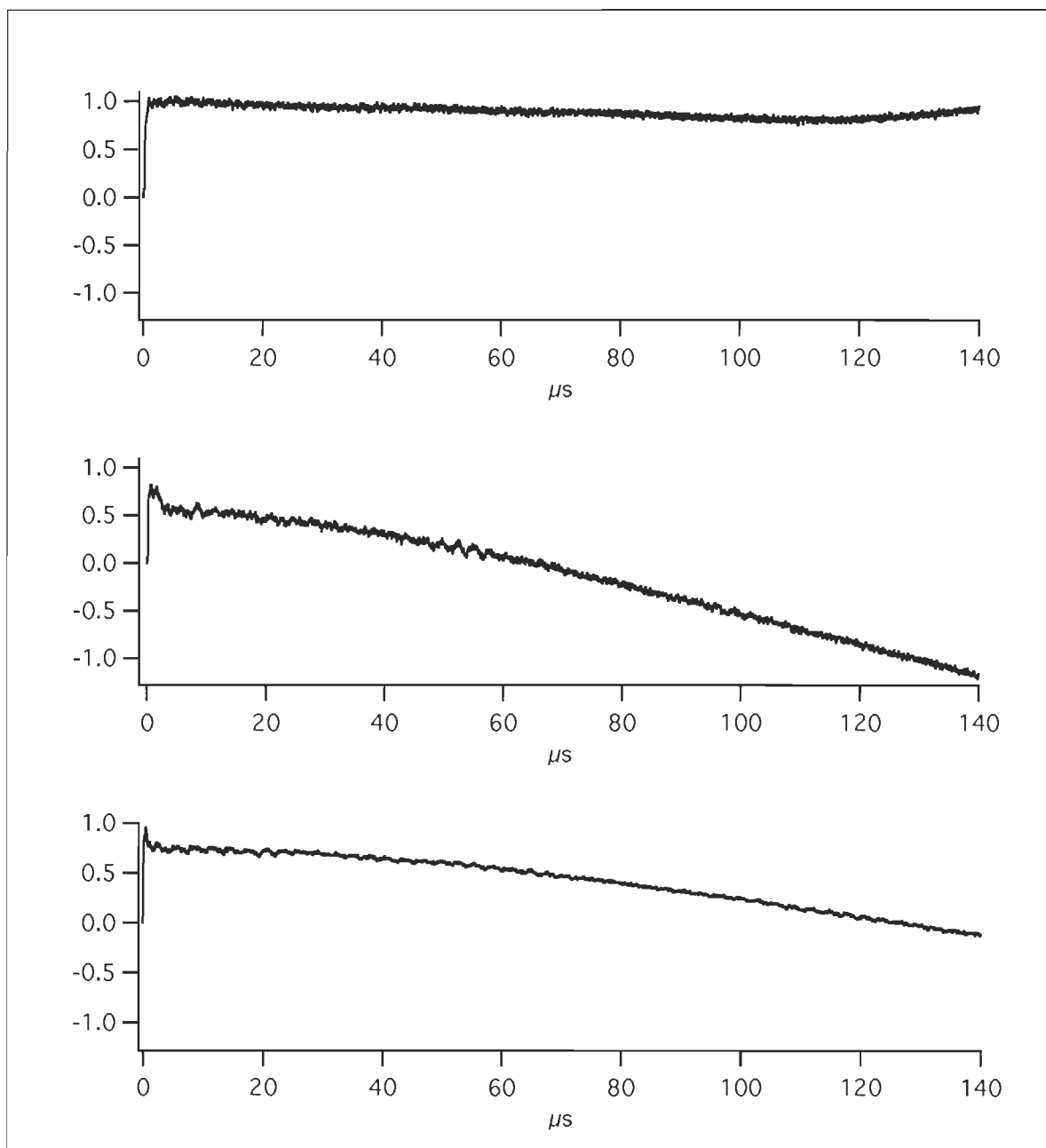
all these small but omnipresent discrepancies between signals of that kind, deconvolution is needed.

But first, the reliability of the deconvolution process with such signals must be established. After each acquisition of a signal from the reference compound, the cell must be emptied and reloaded with the photoactive sample. Such a refilling process inevitably slightly perturbs the system. One must ascertain that these perturbations, in addition to the noise on signals, will not spoil completely the results. To verify that the perturbations due to the refilling process do not unduly alter the results, all the reference signals have been arbitrarily grouped two by two. Since the signals in a group are both reference signals, they are identical except for eventual temperature fluctuations, refilling perturbations and noise. The result of deconvolution of two identical signals is a Dirac delta function whose integral is a Heaviside function.

The upper graph of Fig. 13 is the average of the mutual deconvolution of the signals in each of 6 groups of reference and photoactive signals. (The signals of the photoactive sample too have been grouped two by two.) It departs from a Heaviside function, the ideal result, by 13% in the  $\chi^2$  sense. (The same result at 22.5°C, not shown here, is similar though slightly better with a departure of only 10%.) It is not a perfect Heaviside function, nevertheless it appears sufficiently good to proceed to further analysis.

The curve of the second graph of Fig. 13 is the average of the deconvolution of 6 active sample signals by 6 reference compound signals, at 11.5°C. By comparison with the curve of the upper graph, even before the discrimination in temperature, three aspects are evident. First; at about  $t \approx 0$ , the initial step is smaller than that of the reference curve. It reflects the fact that the initial volume expansion due to immediate conversion of light into heat is less with the photoactive sample. Second; after the initial step, the overall shape of the curve is a

**Fig. 13.** First graph: average of deconvolutions of reference signals by reference signals at 11.5°C. Second graph: average of deconvolutions of purple membranes signals by reference signals at 11.5°C. Third graph: average of deconvolutions of purple membranes signals by reference signals at 22.5°C. All graphs are at the same scale.





long slope. It may be due partially to a decrease in volume of the sample and very likely too, as discussed earlier, to the non-uniformity in the acoustic sources of the signals participating in the deconvolution. Third; at about  $t \approx 2.5 \mu\text{s}$ , there is an obvious and relatively abrupt negative step which accounts for approximately 25% of the maximum of the function, indicative of a contraction in the sample.

The curve of the third graph of Fig. 13 is similar to, but slightly different from the curve of the second graph. It is due to the fact that the thermal expansion coefficient of water is larger at  $22.5^\circ\text{C}$  than at  $11.5^\circ\text{C}$  while the conformational volume changes of purple membranes do not vary, or at least exhibit much small variation than the volume changes of water, for the same change in temperature. Hence, the effect of purple membrane relatively to the effect of water in the sample is much smaller. Consequently, the curve of the third graph will tend to behave more like a Heaviside function than the curve of the second graph, though, the effect of purple membranes remains visible.

Since the experiments have been performed at two different temperatures, it becomes possible to discriminate between the thermal volume changes and the conformational volume changes in the photoactive sample. Thereafter, we develop the needed equations, applicable to the functions obtained by the deconvolution, to perform such a discrimination.

## Temperature discrimination

Suppose a function  $w(t)$ , ( $w(t) = 0$  for  $t < 0$ ) which represents any heating power applied to the sample and  $\epsilon(t)$  the total heat energy released in the sample up to time  $t$ . Evidently we then have:

$$\epsilon(t) = \int_{-\infty}^t w(t) dt, [\text{J}] \quad (52)$$

Let  $V_{0T}$  be the volume of the sample for  $t < 0$ ,  $V_T(t)$  the volume of the sample at time  $t$ ,  $\Delta V_T(t)$  and  $\Delta T_T(t)$  the changes in volume and temperature respectively between time  $t = 0$  and time  $t$ . The subscript  $T$  indicates the temperature at which the experiment is done. Thus:

$$V_T(t) = V_{0T} + \Delta V_T(t), [\text{m}^3] \quad (53)$$

The thermal expansion coefficient at temperature  $T$  is defined as:

$$\alpha_T = \frac{1}{V} \frac{d}{dT} V_T(t), [\text{K}^{-1}] \quad (54)$$

If  $\Delta V_T$  and  $\Delta T$  are small enough, we obtain from eq. (54):

$$\Delta V_T(t) = \alpha_T V_T(t) \Delta T(t) \quad (55)$$

The relation between heat energy and changes in temperature is:

$$\epsilon(t) = C_T \Delta T(t) \rho_T V(t) \quad (56)$$

in which  $C_T$  and  $\rho_T$  are the specific heat and the density at temperature  $T$ . Combining eq. (55) and (56) gives:

$$\Delta V_T(t) = \frac{\alpha_T}{C_T \rho_T} \varepsilon(t) \quad (57)$$

The source of the signal (originating from an acoustic wave in the cell converted through all elements in the chain) is not  $V_T(t)$  nor  $\Delta V_T(t)$ , it is instead the derivative of  $V_T(t)$ . Differentiating eq. (53) and substituting eq. (57) in the result gives:

$$\frac{d}{dt} V_T(t) = \frac{d}{dt} (V_{0T} + \Delta V_T(t)) = \frac{d}{dt} \Delta V_T(t) = \frac{d}{dt} \left( \frac{\alpha_T}{C_T \rho_T} \varepsilon(t) \right) = \frac{\alpha_T}{C_T \rho_T} w(t) \quad (58)$$

Let us denote  $w_l(t)$ , the power of the laser applied to the sample. The laser pulse being very short relative to the characteristic time of the system, it is considered as a Dirac delta function whose amplitude depends on the energy ( $\varepsilon_0$ ) in the pulse. So:

$$w_l(t) = \varepsilon_0 \delta(t) , \text{ [J/s]} \quad (59)$$

Thus, the function  $\varepsilon_l$ , the energy delivered into the sample by the laser is:

$$\varepsilon_l(t) = \int_{-\infty}^t w_l(t) dt = \int_{-\infty}^t \varepsilon_0 \delta(t) dt = \varepsilon_0 H(t) \quad (60)$$

where  $H(t)$  is the Heaviside unit step function. If the sample is a reference solution that almost instantly converts light energy into heat, then eq. (58), when  $w_l(t)$  is in place of  $w(t)$ , becomes:

$$\frac{d}{dt}V_T(t) = \frac{\alpha_T \varepsilon_0}{C_T \rho_T} \delta(t) , \left[ \frac{\text{m}^3}{\text{s}} \right] \quad (61)$$

If the derivative of  $V_T(t)$  given by eq. (61) was a pure Dirac delta function, with its integral equal to unity, then its effect on the system would be  $I_T(t)$ , the characteristic function at temperature  $T$ , which is necessary to compute deconvolutions. Since it is not exactly the case, because the integral from  $-\infty$  to  $\infty$  of eq. [61] gives  $(\alpha_T \varepsilon_0)/(C_T \rho_T)$  instead of unity (which would have been the case if the coefficient of the Dirac function in eq. [61] would have been unity), let us denote by  $U_T(t)$  the response of the system to the function given by eq. (61). Since the system is linear, which implies that  $a \times \text{causes}$  produces  $a \times \text{effects}$ , we immediately have:

$$U_T(t) = \frac{\alpha_T \varepsilon_0}{C_T \rho_T} I_T(t) \quad (62)$$

Considering now the case of photoactive purple membranes, which store a part of the energy from the laser pulse and release it later in the form of heat or conformational changes, we define the following function:

$$E_T(t) = \frac{d}{dt}V_T(t) , \left[ \frac{\text{m}^3}{\text{s}} \right] \quad (63)$$

where  $E_T(t)$  includes a thermal and a conformational component. Let us denote by  $E_{eT}(t)$  the thermal component that depends on the temperature and by  $E_c(t)$  the conformational component which will be considered as independent of the temperature. We directly have:

$$E_T(t) = E_{eT}(t) + E_c(t) \quad (64)$$

The thermal power released by the photoactive sample versus time is denoted  $a(t)$ . We suppose that  $a(t)$  (but not its associated volume change) is independent of the temperature. Since  $a(t)$  does not contribute to the conformational component, then from eq. (58), with  $a(t)$  in place of  $w(t)$ , and eq. (63) and (64) we obtain:

$$E_{eT}(t) = \frac{\alpha_T}{C_T \rho_T} a(t) \quad (65)$$

If we repeat the experiment at two different temperatures  $T_1$  and  $T_2$  we obtain:

$$\begin{aligned} E_{eT_1}(t) &= k_1 a(t) \\ E_{eT_2}(t) &= k_2 a(t) \end{aligned} \quad (66)$$

in which:

$$k_i = \frac{\alpha_{T_i}}{C_{T_i} \rho_{T_i}}, \left[ \frac{\text{m}^3}{\text{J}} \right] \quad (67)$$

From eq. (66) we draw:

$$\frac{E_{eT_1}}{k_1} = \frac{E_{eT_2}}{k_2} \quad (68)$$

Denoting by  $D_{T_i}(t)$  the response of the system to the function  $E_{T_i}(t)$  we have by definition of convolution:

$$D_{T_i}(t) = I_{T_i}(t) \bullet E_{T_i}(t) \quad (69)$$

and then, following eq. (47):

$$E_{T_i}(t) = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{D_{T_i}(t)\}}{\mathcal{F}\{I_{T_i}(t)\}} \right\} \quad (70)$$

Bringing eq. (62) and (67) in eq. (70), we obtain:

$$E_{T_i}(t) = k_i \varepsilon_0 A_i(t) \quad (71)$$

where:

$$A_i(t) = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{D_{T_i}(t)\}}{\mathcal{F}\{U_{T_i}(t)\}} \right\} \quad (72)$$

Eq. (64) expressed in terms of  $T_1$  and  $T_2$  gives the following system:

$$\begin{aligned} E_{T_1}(t) &= E_{eT_1}(t) + E_c(t) \\ E_{T_2}(t) &= E_{eT_2}(t) + E_c(t) \end{aligned} \quad (73)$$

Eliminating  $E_c(t)$  in eq. (73) and bringing eq. (68) in the result gives:

$$E_{T_1} - E_{T_2} = \left( \frac{k_1 - k_2}{k_1} \right) E_{eT_1} \quad (74)$$

Expressing eq. (71) and (72) in terms of  $T_1$  and  $T_2$ , bringing the result in eq. (74) and rearranging the terms gives:

$$E_{eT_1} = \varepsilon_0 \left( \frac{k_1}{k_1 - k_2} \right) (k_1 A_1 - k_2 A_2), \left[ \frac{\text{m}^3}{\text{s}} \right] \quad (75)$$

With eq. (71), (72) and (75) in (73) to solve the system, we get:

$$E_c = \epsilon_0 \left( \frac{k_1 k_2}{k_1 - k_2} \right) (A_2 - A_1), \left[ \frac{\text{m}^3}{\text{s}} \right] \quad (76)$$

Integrating eq. (75) and (76) gives the volume changes due to thermal release at  $T_1$ , ( $\Delta V_{T_1}(t)$ ) and the volume changes due to conformational modifications ( $\Delta V_c(t)$ ):

$$\Delta V_{T_1}(t) = \int_{-\infty}^t E_{eT_1} dt = \epsilon_0 \left( \frac{k_1}{k_1 - k_2} \right) \left( k_1 \int_{-\infty}^t A_1 dt - k_2 \int_{-\infty}^t A_2 dt \right), [\text{m}^3] \quad (77)$$

$$\Delta V_c(t) = \int_{-\infty}^t E_c dt = \epsilon_0 \left( \frac{k_1 k_2}{k_1 - k_2} \right) \left( \int_{-\infty}^t A_2 dt - \int_{-\infty}^t A_1 dt \right), [\text{m}^3] \quad (78)$$

Choosing  $T_1 = 11.5^\circ\text{C}$  and  $T_2 = 22.5^\circ\text{C}$ , we obtain  $k_1 = 25.8 \times 10^{-12} \text{ m}^3/\text{J}$  and  $k_2 = 55.9 \times 10^{-12} \text{ m}^3/\text{J}$ . The energy absorbed by the sample for each laser pulse is  $30 \pm 10 \text{ mJ}$  without mirror. With a mirror we obtain  $\epsilon_0 = 37 \pm 12 \text{ mJ}$ .  $A_1$  and  $A_2$  are the functions resulting from the deconvolutions of the photoactive sample signals by the reference compound signal at  $11.5^\circ\text{C}$  and  $22.5^\circ\text{C}$  respectively. With these values and functions, eq. (77) gives the curve 1 of the upper graph of Fig. 14. It is the thermal volume change of the active sample at  $11.5^\circ\text{C}$ . Eq. (78), with the same parameters and functions gives the curve 4. It is the conformational volume change of the active sample.

If we repeat the computations with the same parameters but, this time, with  $A_1$  and  $A_2$  being the result of the deconvolution of reference compound signal by another reference compound signal at  $11.5^\circ\text{C}$  and  $22.5^\circ\text{C}$  respectively, we obtain the gray curves (2 and 3) of the upper graph of Fig. 14. Curve 2 is the thermal volume change of the reference compound so-

lution at 11.5°C. Curve 3 is the conformational volume change of the reference compound solution.

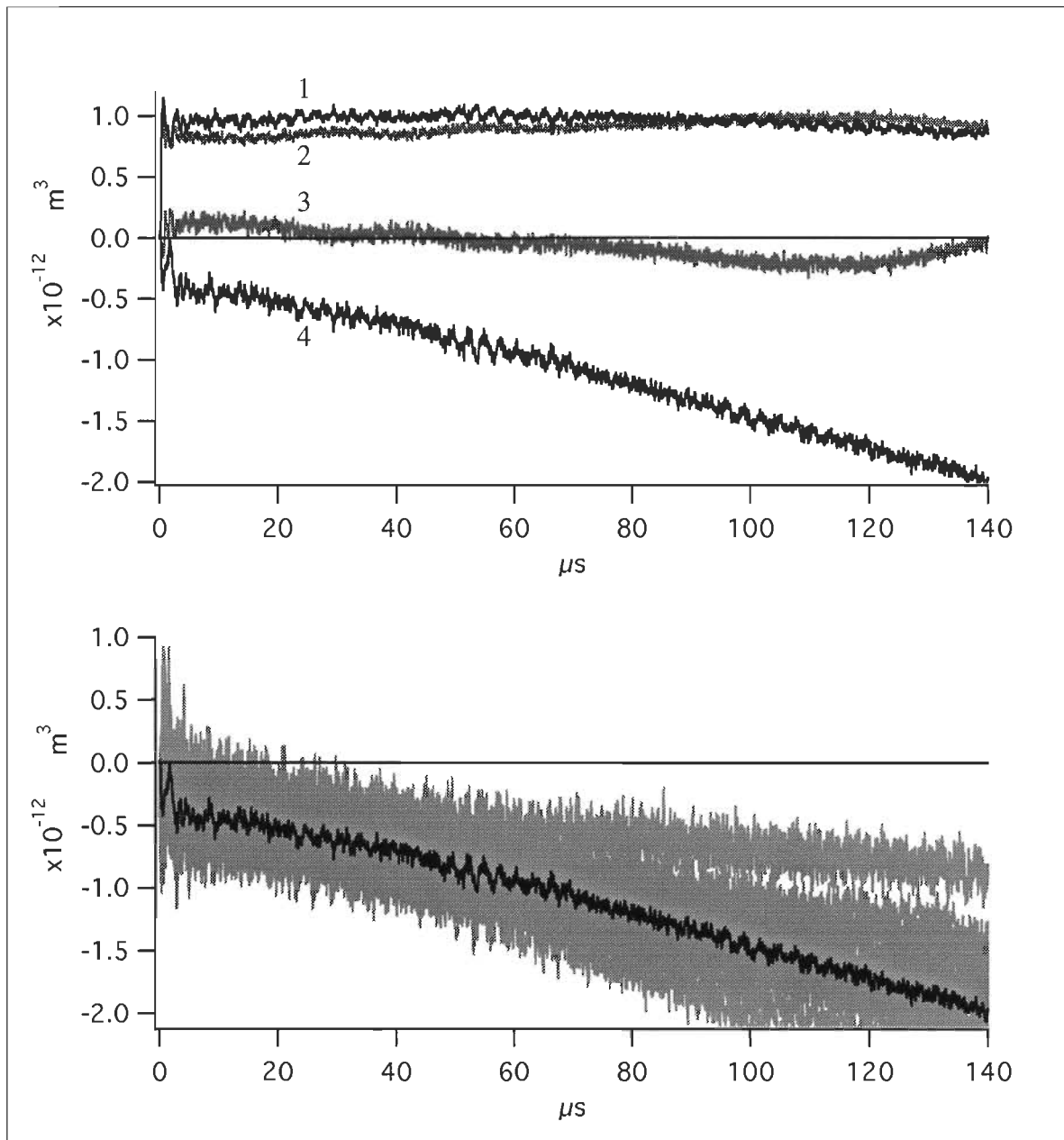
About these results we can say, first, that curve 1 and 2 of the upper graph of Fig. 14 should be almost equal. In fact, curve 2 should correspond to the volume change induced by the a laser pulse of an energy of 37 mJ completely transformed into heat. Curve 1 should be the same but slightly lower. The difference corresponding to the energy stored by the bacteriorhodopsin in the br->K transition of its photocycle. If all br in the excited volume was in the K state (with an energy content of 160 kJ/mole [27]) at the end of the laser flash, then, at an absorbance of 0.3, with a molar absorption coefficient of 45000 at 532 nm, it would mean that at most 168  $\mu$ J of the 37 mJ of the laser pulse were stored. It would mean a difference in thermal volume change of 0.5% (invisible on the graph). The volume swept by the laser beam in the cell is  $159 \times 10^{-9} \text{ m}^3$ . This volume of water, at 11.5°C, heated by 37 mJ, gives a volume change of  $0.95 \times 10^{-12} \text{ m}^3$ . The curves 1 and 2 are, as expected, approximate Heaviside functions of average amplitude of  $0.96 \times 10^{-12} \text{ m}^3$  and  $0.91 \times 10^{-12} \text{ m}^3$  respectively. These three values almost agree. The discrepancy between curve 1 and 2 gives an idea of the accuracy of the whole process.

Second; the curve 3 of the upper graph of Fig. 14 is the conformational volume change of the reference compound. It should be zero everywhere. It is only approximately zero. Here too it gives an idea of the accuracy of the whole process.

Third; the curve 4 represents the conformational volume changes of the photoactive sample. Note that all the photoactive signals have been obtained from sample coming from the same solution of purple membranes in pure water. Results at 11.5°C and at 22.5°C have been obtained during subsequent measurement sessions at the laboratory. The curve 4 is, in



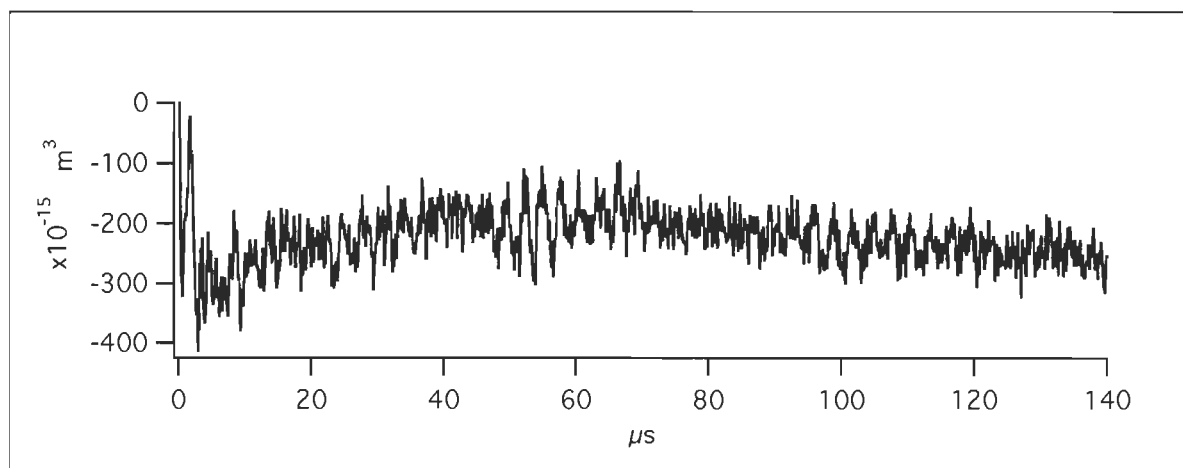
**Fig. 14.** Upper graph: Results of temperature discrimination of components of signals issued from deconvolutions. Curve 1 is the thermal expansion of water in the active sample. Curve 2 is the thermal expansion of water in the reference compound solution. Curve 3 is the conformational volume changes of water in the reference compound solution. Curve 4 is the conformational volume changes of br in the active solution. Lower graph: the black curve is the curve 4 of upper graph. The gray background is the stacking of all the functions which gave, when averaged, the black curve.



fact, as for the other curves, the average of all the results from all the available deconvolutions. This curve is reproduced in black in the lower graph of the Fig. 14. The gray background of this graph is the stacking of all the fifteen results functions that have been averaged to give the final result. It shows the dispersion of the results and another idea of the accuracy of the whole process.

Furthermore, the overall shape of the curve 4 is a long negative slope. As discussed earlier, it may be due partly to a decrease in volume of the sample and very likely too, to the non-uniformity in the acoustic sources of the signals participating in the deconvolution. If we suppose, which is very unlikely to be exact, that the component due to this non-uniformity is the straight line which would be the best fit to the curve 4 between  $2.5 \mu\text{s}$  and  $140 \mu\text{s}$ , and if we subtract this line from curve 4 and pull back the first point of the function to zero, we obtain the curve of Fig. 15. It is an interesting curve but drawing more conclusions from it at this time would be hazardous. A completely new, different and more accurate system should be built to clear up this question.

**Fig. 15.** Conformational volume changes of br, arbitrarily corrected. (See text)



Nevertheless the curve 4 of Fig. 14, as the curve of the lower graph of Fig. 13, show obvious and relatively sharp negative steps at about  $t \approx 2.5 \mu\text{s}$ . It represents a contraction of  $420 \times 10^{-15} \text{ m}^3$ . The accuracy of the amplitude of this contraction and the precise time of its occurrence may be arguable but its mere existence is hardly contestable since nothing similar is present on any of the reference curves which have been obtained through the same physical, mathematical and numerical treatments.

It is admitted that during the few nanoseconds of the laser flash a photoequilibrium is established between br and its K photointermediate [27, 30]. At the end of the flash, only the molecules in the K state will accomplish the photocycle. Let us nevertheless suppose that all the molecules in the excited volume are in the K state at the end of the laser flash. Then the contraction of  $420 \times 10^{-15} \text{ m}^3$  at  $t \approx 2.5 \mu\text{s}$  means a contraction of  $665 \times 10^{-30} \text{ m}^3/\text{molecule}$ . It represents a volume change of about 1% of the volume of the molecule, or the volume of about 20 molecules of water. It may seems small but according to values found in the literature it is quite large. In [45] a volume expansion of  $100 \times 10^{-30} \text{ m}^3/\text{molecule}$  is reported instead of a contraction. The authors of [33] have obtained a contraction of  $3.7 \times 10^{-30} \text{ m}^3/\text{molecule}$ . In the same paper, a comparison between different results of volume changes in the microsecond time scale is presented. It goes from  $-20 \times 10^{-30} \text{ m}^3/\text{molecule}$  to  $100 \times 10^{-30} \text{ m}^3/\text{molecule}$ . All these results are summarized in Table 15. As we can see, there is still a controversy on that point. Nevertheless, our result of  $-665 \times 10^{-30} \text{ m}^3/\text{molecule}$  appears as an outlier. There are two possible reasons for such a result.

First; the excited volume of the sample may be much larger than the cylinder in the cell with the diameter of the laser beam (4.5 mm). The walls of the cell are the imperfect faces of a square glass tube. The laser beam is distorted and diffused when it goes through these walls

**Table 15.** Different results for the conformational volume change of the bacteriorhodopsin near the K to L stage of the photocycle. These changes occur in the microsecond scale (1-3  $\mu$ s).

Schulenberg... ref. [45]	$+100 \times 10^{-30} \text{ m}^3 / \text{molecule}$
Zhang... ref. [33]	$- 3.7 \times 10^{-30} \text{ m}^3 / \text{molecule}$
Varo...ref. [33]	$- 20 \times 10^{-30} \text{ m}^3 / \text{molecule}$
Present results	$- 665 \times 10^{-30} \text{ m}^3 / \text{molecule}$

and through the sample. This perturbed beam is returned back into the cell with the mirror. Hence, it is very likely that the excited volume of sample is larger than a cylinder of 4.5 mm in diameter, resulting in a greater number of excited molecules. For an excited volume of 1  $\text{cm}^3$ , the measured volume change would be  $-106 \times 10^{-30} \text{ m}^3/\text{molecule}$ , a result comparable in absolute amplitude to that of ref. [45] but with a different sign.

Second; we have used an energy density of  $\sim 10^{18} \text{ photons/cm}^2$  which is  $\sim 10^2$  and  $\sim 10^4$  times more than the density used by the authors of ref. [45] and ref. [33] respectively. As discussed in [27] and [33], such a high density could alter and even lead to overestimate the results. Nevertheless, with our rudimentary system, our main concern was more to test the possibility of using Fourier transform for deconvolution of TRPAS signals than to obtain perfectly accurate and definitive results. Essentially, this objective has been fulfilled.

## Conclusion

One of the main purpose of this paper is to discuss the pertinence to use Fourier transform for deconvolution of real, digitized and noisy signals in general and, more particularly, those from TRPAS systems. It has been shown that an almost mandatory prerequisite to obtain correct deconvolution results is the use of a highly accurate numerical Fourier transform method. Without such a method, the results may be greatly distorted, even with noiseless signals.

Another important obstacle to a wide use of Fourier transform in deconvolution is the so-called noise generation due to the division in the transformed domain. In fact, as discussed, there is not really a “generation” of noise; the noisy result of deconvolution is the necessary behavior of a good deconvolution system for noisy input signals. It has been shown that Fourier transform is usable and perform very well for deconvolution if the noise in the initial signals is sufficiently reduced by any means which does not destroy the seeked information. A very good way to do so is simply to average a sufficient number of initial signals.

A real test with a TRPAS system submitted to multiple laser pulse, recorded by photodiode for comparisons, have shown that a proper deconvolution method can produce correct results, in amplitude as well as in time. This same test reveals that the two time-related limitations in TRPAS experiments, the smallest time resolution linked to the diameter of the laser beam and the maximum duration of measurements, limited by the reflections of signals, can be removed with proper deconvolution. Another difficulty with TRPAS systems, not related to but exposed by deconvolution, is that of the amplitude dissymmetry between the sources of acoustic signals in reference and photoactive signal. We have discussed the effect of this dissymmetry in the case of our particular system. This problem should be addressed to obtain correct results with any TRPAS systems.

Tests performed with purple membranes have shown that it is possible to obtain significant and plausible results. Once the equations for temperature discrimination were established and applied to results of deconvolution, signals prove to be clean enough to allow estimation of conformational volume changes of bacteriorhodopsin.

**Bibliography**

19. J.E. RUDZKI, J.L. GOODMAN and K.S. PETERS. *J. Am. Chem. Soc.* **107**, 7849 (1985).
20. D.A. HUTCHINS. *Can. J. Phys.* **64**, 1247 (1986).
21. L.A. MELTON, T. NI and Q. LU. *Rev. Sci. Instrum.* **60**(10), 3217 (1989).
22. S.E. BRASLAVSKY and K. HEIHOFF. *In Handbook of Organic Photochemistry*. J.C. SCAIANO. CRC Press. 1989. p. 327.
23. K.S. PETERS, T. WATSON and K. MARR. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 343 (1991).
24. S.E. BRASLAVSKY and G.E. HEIBEL. *Chem. Rev.* **92**(6), 1381 (1992).
25. P.R. CRIPPA, A. VECLI and C. VIAPPIANI. *J. Photochem. Photobiol.* **24**, 3 (1994).
26. K. MARR and K.S. PETERS. *Biochemistry.* **30**(5), 1254 (1991).
27. M. ROHR, W. GÄRTNER, G. SCHWEITZER, A.R. HOLZWARTH and S.E. BRASLAVSKY. *J. Phys. Chem.* **96**(14), 6055 (1992).
28. D. BRUCE and O. SALEHIAN. *Biochimica et Biophysica acta.* **1100**, 242 (1992).
29. K.S. PETERS, T. WATSON and T. LOGAN. *J. Am. Chem. Soc.* **114**(11), 4276 (1992).
30. P.J. SCHULENBERG, M. ROHR, W. GÄRTNER and S.E. BRASLAVSKY. *Biophysical Journal.* **66**, 838 (1994).
31. J-L. HABIB JIWAN, A. K. CHIBISOV and S.E. BRASLAVSKY. *J. Phys. Chem.* **99**, 10246 (1995).
32. M.E. van BREDERODE, T. GENSCH, W.D. HOFF, K. J. HELLINGWERF and S.E. BRASLAVSKY. *Biophysical Journal.* **68**, 1101 (1995).
33. D. ZHANG and D. MAUZERALL. *Biophysical Journal.* **71**, 381 (1996).
34. C.K.N. PATEL and A.C. TAM. *Reviews of Modern Physics.* **53**(3), 517 (1981).
35. A.C. TAM. *Reviews of Modern Physics.* **58**(2), 381 (1986).
36. J. SCHUTTE. *Rev. Sci. Instrum.* **52**(3), 400 (1981).
37. S. MAKINEN, *Rev. Sci. Instrum.* **53**(5), 627 (1982).

38. S. SORELLA and S.K. GHOSH. *Rev. Sci. Instrum.* **55**(8), 1348 (1984).
39. M. FROEYEN and L. HELLEMANS. *Rev. Sci. Instrum.* **56**(12), 2325 (1985).
40. N. BEAUDOIN. *Canadian Journal of Physics.* **76**(9), 659 (1998).
41. J. C. ANDRE, L. M. VINCENT, D. O'CONNOR and W. R. WARE. *J. of Phys. Chem.* **83**(17), 2285 (1979).
42. J. N. DEMAS. *Excited States Lifetime Measurements.* Academic Press. 1983. p. 148.
43. P. J. ROUSSEEUW. *Handbook of Statistical Methods for Engineers and Scientists.* Edited by H. M. WADSWORTH. McGraw-Hill. 1990. p. 16.1.
44. S. E. BRASLAVSKY. *The Spectrum.* **7**(4),10 (1994).
45. P.J. SCHULENBERG, W. GÄRTNER and S.E. BRASLAVSKY. *J. Phys. Chem.* **99**(23), 9617 (1995).



## **CONCLUSION**

## **Synoptique**

Le premier point à mentionner en conclusion est sans doute le fait que nos objectifs furent atteints. Nous voulions voir ces “tremblements de protéines”. Nous les avons vus! Évidemment la route fut longue et parsemée d’embûches de toutes sortes. Mais à force de temps et surtout d’une confiance indéfectible dans une certaine philosophie du travail et de l’approche des questions scientifiques, les problèmes furent, un à un, résolus.

Nous avons toujours l’impression, lors de la confection d’un rapport de recherche, que ce dont nous devons et pouvons parler ne représente qu’une faible partie de tout ce qu’il a fallu faire pour atteindre quelques résultats intéressants. C’est, d’une part, pour atténuer un peu cette impression et, d’autre part, pour mentionner ou préciser certains aspects du travail que nous allons, avant de conclure formellement, passer en revue quelques éléments qui ne pouvaient trouver leur place dans le cadre d’articles scientifiques, mais qui n’en demeurent pas moins importants pour la compréhension de la problématique, ou tout simplement, pour la quiétude d’esprit du lecteur. L’ordre de présentation de ces éléments suivra l’ordre logique ou chronologique de la génération d’un résultat final, à travers toute la chaîne des composants du système SPART, et ce, à partir de l’émission de l’impulsion laser.

## Le laser et les impulsions multiples

Le premier élément de la chaîne est bien sûr notre bon vieux laser YAG (Yttrium Aluminium Garnet). Il s'agit d'un laser des premières générations de ce type. Il exhibe de ce fait certaines qualités et certains défauts reliés à ses origines. C'est une grosse machine encombrante, pas très puissante et capricieuse. Par contre, tous ses éléments sont facilement accessibles, démontables, et réparables. Sous plusieurs aspects, elle semble incomplète. Les montures d'ajustements fins sont pratiquement inopérables. Les cristaux sensibles à la température et à l'humidité ne sont pas scellés. Il a fallu lui construire un système de maintien de la température. On ne compte plus les fois où il a fallu démonter, réaligner, remonter et réajuster cette machine pour la garder en vie. Bien souvent, pour des raisons de coûts ou de disponibilité, il a même fallu construire de toutes pièces les outils pour ce faire: système de photodiodes rapides, amplificateurs, absorbeurs, montures... Cependant, cela nous a amenés à étudier et comprendre assez à fond les principes de fonctionnement des lasers en général, et de ce type de laser en particulier.

Un élément important dans la vérification de la validité des résultats produits par l'ensemble du système SPART fut la possibilité de générer une impulsion laser multiple servant à mimer la réponse d'une solution photoactive. Revoyons donc brièvement comment fonctionnait notre laser afin d'être en mesure de mieux saisir la façon dont nous l'avons modifié pour ce faire.

Une puissante lampe à décharge illumine pendant quelques centaines de microsecondes la tige lasante, un cristal YAG. Un interrupteur optique (Q-switch) placé dans la cavité résonnante, i.e. entre les deux miroirs, empêche l'effet laser de se produire et la tige de se décharger. Après un intervalle de temps précis, déterminé lors de l'ajustement du laser, la charge d'électrons pompés dans la tige YAG est à son maximum. Alors, l'interrupteur est brusque-

ment ouvert et l'impulsion laser de quelque 15 nanosecondes émerge avec puissance. Cette impulsion est un paquet d'ondes électromagnétiques dont la longueur d'onde centrale est de 1064 nm. Le faisceau (on peut parler de faisceau, car la durée de l'impulsion donne un javelot de lumière de quelque 5 m de long) est acheminé sur un cristal transparent qu'il traverse. Lorsque l'intensité du faisceau et donc le champ électrique en son sein sont assez grands, l'état d'excitation des électrons du cristal atteint des régions non-linéaire. C'est un fait mathématique que tout signal, si monochromatique soit-il, génère des harmoniques, lorsque transmis par un élément non-linéaire. Conséquemment, au sortir de ce cristal, le faisceau laser est constitué d'une superposition de la fondamentale à 1064 nm et de toutes les harmoniques, lesquelles évidemment décroissent d'intensité avec la longueur d'onde. Par une série de prismes et de collimateurs nous ne conservons que la composante verte à 532 nm; une longueur d'onde d'un beau vert clair, adéquate pour nos expériences.

L'interrupteur optique est constitué de deux polariseurs dont les plans de polarisation sont parallèles. Un de ces polariseurs est constitué d'un cristal dont le plan de polarisation tourne sous l'effet d'un champ électrique de quelques milliers de volts par cm. Un ajustement très méticuleux et très capricieux de cette tension (car dépendant aussi de la température et de l'orientation du cristal) amène le plan de polarisation du cristal, perpendiculaire à celui du polariseur fixe. L'interrupteur optique est alors fermé. Un circuit électronique de synchronisation, lui aussi soigneusement ajusté, produit au bon moment une impulsion qui porte en avalanche une chaîne de transistors, laquelle est montée de façon à court-circuiter les électrodes de polarisation du cristal. La tension sur le cristal devient momentanément nulle, son plan de polarisation se rétablit, l'interrupteur optique est ouvert, l'effet laser se produit.

C'est à ce point que s'introduit le générateur d'impulsions multiples mentionné dans le deuxième article. Nous avons construit un circuit électronique qui intercepte le signal de mise en avalanche des transistors et l'achemine à l'entrée d'une cascade de 3 circuits intégrés,

montés en monostable. Chacun de ces circuits intégrés, après un délai ajustable, génère à son tour une impulsion. Les impulsions de tous ces monostables sont ensuite conduites à un mélangeur, lequel donne enfin le train d'impulsions requis. C'est ce train d'impulsions qui est retourné à la chaîne de transistors de l'interrupteur optique. Ainsi, pendant une seule décharge de la lampe de pompage, on a pu ouvrir 3 fois l'interrupteur optique.

On obtient ainsi, pendant un seul tir, 3 impulsions laser, séparées chacune de quelques microsecondes à quelques dizaines de microsecondes. L'intensité de la lampe à décharge du laser ayant un profil à peu près de forme  $ate^{-bt}$ ,  $a, b > 0$ , il est possible, en positionnant les 3 impulsions adéquatement dans le temps pendant l'éclair de la lampe, de faire varier les amplitudes des 3 impulsions laser résultantes. Évidemment, les amplitudes des impulsions laser et les temps où elles surviennent ne sont pas indépendants. Cependant, il nous restait toujours assez de souplesse pour produire des signaux adéquats pour les fins auxquelles ils étaient destinés.

Un autre petit circuit, fait maison, constitué d'éléments d'accord d'impédance et d'une batterie de photodiodes rapides (MRD-500) fortement polarisées en inverse pour en réduire au maximum la capacitance et ainsi en augmenter la vitesse, a été utilisé pour "voir" sur oscilloscope les impulsions laser. Nous avons donc construit tout ce qu'il fallait pour contrôler la véracité des résultats de notre système SPART, incluant les déconvolutions de ses signaux réels et bruyants.

## **La cellule et le froid**

Les impulsions laser requises étant disponibles, acheminons-nous maintenant vers leur destination: la cellule munie de son capteur, lequel est constitué d'un film de plastique piézo-électrique pressé entre le fond de la cellule et un cylindre plein fait de plomb. Le plomb sert à atténuer les oscillations parasites (ringing). Nous avons suffisamment parlé en introduction et dans le deuxième article des contraintes présidant à la mise au point d'une cellule fonctionnelle. Nous ne retiendrons ici que quelques points descriptifs et quelques faits cocasses.

Pour des raisons déjà mentionnées, notre cellule, constituée d'un tube de verre carré, fermé au bas d'un disque de quartz collé à l'époxy, avait plus d'un mètre de long. Il fallait pouvoir vidanger fréquemment cette cellule sans la bouger et sans presque lui toucher. Un système de trompe à eau et un petit piston carré de teflon permettait d'effectuer la tâche avec beaucoup de délicatesse.

Notre bain thermostaté étant trop vibrant et pas assez puissant pour une si longue cellule, l'ajustement de la température, nécessaire pour la discrimination des composantes thermique et conformationnelles des signaux, fut une constante peine. Tout le laboratoire devait être amené et maintenu à la température requise. Les mesures à 22°C, bien que difficiles à stabiliser furent tout de même praticables. Par contre les mesures à basse température ne furent atteintes qu'en démontant les fenêtres et en mettant à profit notre merveilleux hiver canadien et/ou québécois. Après quelques semaines de mesures acceptables, il est apparu que ce n'étaient pas les meilleures conditions pour le matériel, non plus que pour l'expérimentateur ou ses relations sociales...

## La solution de référence

Nous avons mentionné dans le deuxième article que notre solution de référence était de l'eau distillée, amenée à la même absorbance que la solution de membranes pourpres, par l'ajout d'encre noire. Dans la littérature, sauf en un endroit où aussi de l'encre noire fut utilisée, l'usage de solutions de chlorure de cuivre ou de bromocrésol pourpre semble de mise. L'encre, d'usage beaucoup plus simple et donnant une absorbance beaucoup plus stable, fut adoptée pour nos mesures. Cependant, préalablement, nous fîmes des comparaisons de signaux, produits avec de l'encre, du chlorure de cuivre et du bromocrésol pourpre. Ces signaux demeuraient indiscernables quel que fut le colorant utilisé. De plus, puisqu'une solution d'eau et d'encre ne contient, à la différence d'une solution d'eau et de membranes pourpres, ni lipides ni protéines; il fallait vérifier si les propriétés thermo-acoustiques de notre solution de référence étaient bien les mêmes que celles de la solution active. Pour ce faire nous avons pris deux solutions d'eau et d'encre de même absorbance et avons ajouté à l'une d'elles de l'albumine sérique bovine (une protéine) jusqu'à une concentration équivalente à celle d'une solution de bactériorhodopsine d'absorbance égale à 2. Puisque les absorbances utilisées pour les mesures subséquentes furent de 0.3, cela représentait une marge de sécurité suffisante. Aucune différence entre les signaux produits par ces deux solutions ne fut observée. Ce résultat, conforme à ceux de la littérature, nous a permis de croire que nos solutions actives et de références avaient des propriétés thermo-acoustiques visiblement identiques.

## **La puissance du laser**

Dans les équations, apparaissant dans le deuxième article, ayant pour fins la discrimination en température entre les composantes thermiques et conformationnelles des signaux déconvolués, une quantité importante est présente. Il s'agit de la quantité d'énergie venant du faisceau laser et absorbée par l'échantillon. N'ayant pas de joulemètre adéquat, cette quantité fut déterminée par calorimétrie. Un nombre connu d'impulsions laser furent lancées dans une cuvette de masse connue, contenant une quantité connue de solution de référence. La variation de température fut mesurée. Les mesures de température furent poursuivies suffisamment longtemps après la fin de la salve d'impulsions laser, pour estimer, pendant un même intervalle de temps que pendant la chauffe, les déperditions de chaleur dans le milieu ambiant. Nous avons obtenu, bien qu'entaché d'une erreur assez grande, le résultat cité dans le deuxième article, soit environ 30 mJ. Ce qui est très plausible considérant notre type de laser.



## Logiciels, programmes et machines

Il y a peu à dire qui n'ait déjà été dit concernant les éléments subséquents de la chaîne du système; soient le capteur, amplificateur, oscilloscope, synchronisateur et ordinateur. Nous arrivons donc en bout de chaîne à la manipulation et au traitement des signaux. Quelques mots d'abord sur le choix des logiciels utilisés.

Pour manipuler et tester avec grande précision les équations apparaissant dans les développements du premier article, il fallait un logiciel puissant de calcul symbolique assorti de grandes capacités numériques. Nous avons étudié et testé les logiciels Theorist, Mathematica et Maple. Theorist est d'usage très souple, mais il n'est pas programmable et la précision des calculs est de moins de 20 chiffres significatifs. Mathematica et Maple se comparent assez bien. Cependant les algorithmes internes de Maple nous sont apparus plus robustes, plus puissants et plus rapides que ceux de Mathematica. Maple nous a semblé aussi plus souple et d'une syntaxe plus fraîche. La précision (le nombre de chiffres significatifs utilisés dans les calculs) n'est ajustable qu'au début d'une session de travail, par une commande spéciale, dans Mathematica. Dans Maple, la précision est ajustable n'importe où et n'importe quand; on peut donc accorder toute la précision voulue là où il le faut et revenir à de basses précisions de quelques dizaines de chiffres significatifs pour accélérer les calculs moins critiques. Les capacités de transport des programmes Maple sont aussi très utiles. Nos programmes développés et testés sur un petit ordinateur de table Macintosh pouvaient être exécutés directement sur l'ordinateur central de l'Université, un IBM RISC 6000, lorsque de grandes capacités de mémoire étaient requises. Nous avons donc opté pour Maple. Tous ces logiciels sont interprétés et non compilés; bien que très puissants et d'une utilité incomparable pour les développements mathématiques et leur vérification, ils demeurent cependant trop lents pour les traitements numériques massifs et répétitifs.

Pour l'acquisition des données, le traitement numérique des signaux et la mise en graphique, nous avons opté pour le logiciel Igor de WaveMetrics. C'est un logiciel d'une souplesse et d'une vitesse exceptionnelle. Il est programmable et toutes les fonctions prédéfinies ou ajoutées sont compilées. Nous avons pu, grâce aux "opérations externes" de Igor, réaliser un programme d'acquisition de données pour interfacer l'ordinateur du laboratoire avec les oscilloscopes numériques Nicolet dont nous faisons usage.

Le programme de déconvolution comme tel fut réalisé dans Igor. C'est un programme qui, dans le format de ce texte, fait 55 pages. Nous ne le mettrons donc pas en annexe! Il comporte les algorithmes d'affinement des transformée de Fourier directe et inverse présentés dans le premier article, la procédure de déconvolution par transformée de Fourier ainsi que des fonctions de contrôle et d'affichage graphique des résultats.

Un programme que nous avons développé dans Maple (de Waterloo Maple Software) et donnant tous les coefficients des filtres Savitzky-Golay symétriques ou asymétriques, pour n'importe quel ordre et n'importe quel nombre de points, a fourni le matériel à l'implantation dans le programme Igor de déconvolution d'un filtre SG d'ordre 5 à 41 points. Nous avons développé, aussi dans ce programme, des filtres de type Wiener ainsi que deux algorithmes "maison", pour lesquels nous n'avons pas de noms, basés sur l'usage d'estimateurs robustes, pour contrôler le bruit des signaux. Il est cependant apparu, après moult tests, plus sage et plus sécuritaire, dans le sens de ne pas altérer l'information contenue dans les signaux, de déconvoluer tous nos signaux sans leur appliquer aucun filtre autre que le "moyennage", qui n'est pas, à proprement parler, un filtre.

## **Conclusion générale et perspectives**

Et maintenant, sans reprendre textuellement le contenu des résumés ou conclusions des articles, survolons quand même les principaux points que nous pouvons retenir de l'ensemble de ce travail.

D'abord, comme nous le mentionnions en entrée de conclusion; nous avons vu ce que nous voulions voir! Au delà des aspects qualitatifs, des résultats quantitatifs concernant les changements de volume de la bactériorhodopsine furent obtenus. Il fut même possible, malgré l'aspect un peu fruste des signaux, d'appliquer les équations de séparation en température, spécialement développées pour des fonctions déconvoluées, et d'obtenir des résultats plausibles concernant les composantes thermiques et conformationnelles de la bactériorhodopsine. Il est cependant à peine besoin de dire, considérant l'appareillage relativement rudimentaire et tenant plus du prototype que du système clefs en main, que les résultats obtenus avec la bactériorhodopsine doivent être considérés comme préliminaires. Le but réalisable fut plus de démontrer la faisabilité de la chose que l'obtention de résultats définitifs et absolument incontestables.

Pour cela, il faudra concevoir un système nouveau comprenant bien sûr un laser stable muni d'un système multi-impulsions, d'un joulemètre précis et d'un analyseur de faisceau. Sera requis aussi une cellule redessinée, thermostatée bien sûr, et permettant à la fois d'éviter les réflexions de surface, les bulles et les dissymétries de la source acoustique. Un capteur plus sensible, des amplificateurs à large bandes et un oscilloscope numérique plus rapide et à mémoire plus profonde seront nécessaires. Finalement, un ordinateur rapide du genre powermac G3-400MHz, permettant d'effectuer les déconvolutions presque en temps réel, sera un atout majeur, particulièrement durant les phases d'ajustement et de mises au point. Avec un tel système, il sera, par exemple, possible de faire tout un ensemble d'expériences

du même genre que celles faites en photolyse-éclair, sur la bactériorhodopsine ou la rhodopsine, dans différents environnements et conditions, et d'en tirer de précieuses informations.

La déconvolution par transformée de Fourier des signaux de SPART rendant possible l'observation sur un long domaine de temps a permis de mettre en évidence un problème qui n'est ni traité ni mentionné dans la littérature. Il s'agit de la dissymétrie d'amplitude de la source de l'onde acoustique. Il apparaît évident que les effets de cette dissymétrie ne peuvent être ignorés.

Avec un signal synthétique produisant plusieurs déversements successif d'énergie dans la cellule, il fut clairement montré que deux limitations importantes des systèmes SPART conventionnels pouvaient être levées. D'abord, la résolution en temps d'un SPART à déconvolution par transformée de Fourier peut être beaucoup plus fine que la limite généralement acceptée, déterminée par le temps mis par l'onde acoustique pour parcourir une distance équivalente à celle du diamètre du faisceau laser. Ensuite, nous avons vu que la durée d'observation, traditionnellement restreinte par la première réflexion de l'onde acoustique, peut dépasser de beaucoup cette limite, puisque que la déconvolution par transformée de Fourier peut prendre en compte toutes les complexifications des signaux introduites par toutes les réflexions.

Un point qui ressort fortement aussi, est que pour faire de bonnes déconvolutions par transformée de Fourier, il faut d'abord avoir une transformée de Fourier numérique précise. Il est en effet primordial que la transformée de Fourier numérique puisse s'approcher d'aussi près que nécessaire de ce que serait une transformée analytique de la version continue des signaux. Sans cela, les résultats peuvent être le siège d'importantes distorsions.

Avec un système adéquat de discrimination, appliqué à un ensemble suffisant de résultats, il fut montré que la bête noire des déconvolutions par transformée de Fourier de signaux

bruyants peut être vaincue. Par un procédé aussi simple que le “moyennage” et l’élimination des résultats s’écartant trop de la moyenne, il fut possible d’obtenir d’excellents résultats; même si chaque résultat de déconvolution pris un à un pouvait sembler très imprécis.

Autant par simulations numériques que par signaux physiques réels générés par le système d’impulsions laser multiples, autant pour des fonctions pures que pour des signaux bruyants, on a vu que la déconvolution par transformée de Fourier pouvait rendre des résultats très précis, tant en amplitude qu’en phase. Nous espérons ainsi avoir quelque peu contribué à réhabiliter la déconvolution par transformée de Fourier en tant que méthode de déconvolution, non seulement acceptable, mais aussi, puissante, directe et d’un usage, somme toute, assez simple.

Mentionnons finalement que fut développée une méthode mathématique apparemment tout à fait nouvelle, à la fois unificatrice et modulaire. Cette méthode qui fait l’objet de tout le premier article fut aussi présentée au Colloque Math-Image de l’UQTR en juin 1997 et au Colloque des Sciences Mathématiques du Québec en octobre 97. C’est une méthode générale qui ne se rapporte pas qu’aux déconvolutions mais qui, évidemment, fut avantageusement utilisée pour cela.

C’est une méthode simple, basée sur des éléments mathématiques bien connus. Facile à comprendre et à programmer, elle n’en demeure pas moins très puissante. Elle se compare très avantageusement aux autres méthodes du genre. Par exemple, la précision, ajustable, de ses résultats, croît exponentiellement avec l’ordre du système de calcul. Elle permet d’obtenir, avec grande précision et sans limites de bandes, la transformée de Fourier d’une fonction numérisée. Elle donne les polynômes splines de tout ordre d’une telle fonction. Elle produit aussi les conditions frontières optimales de ces fonctions splines. Elle permet, toujours pour une fonction numérisée, d’en calculer très précisément les dérivées de tout ordre ainsi que

l'intégrale entre n'importe quelles bornes. Sa conception même rend possible l'estimation de l'erreur sur n'importe quel des résultats qui en découlent. Elle peut être mise en oeuvre comme un tout, ou bien certaines de ses parties peuvent en être extraites et utilisées seules pour produire des résultats spécifiques, ou encore comme modules à l'intérieur de d'autres méthodes.

Voilà! Ici se termine ce travail. Je pense que nous pouvons être heureux que nos objectifs aient pu être atteints et peut-être même, dans certains cas, dépassés. Comme toujours la réponse à quelques questions ou la solution de quelques problèmes ouvre de multiples perspectives. Ce peut être la possibilité, déjà mentionnée, de réaliser un système de spectroscopie photoacoustique à résolution temporelle quasi clefs en main, exempt des inévitables carences d'un prototype. Il y a aussi les possibles suites, conséquences, ou ramifications de l'un ou l'autre des points développés dans cette thèse. On pense ici, au delà du développement du système comme tel, à toutes les applications possibles, tant biophysiques que physiques ou chimiques d'un système de spectroscopie photoacoustique à résolution temporelle. Dans le domaine mathématique, la disponibilité d'une transformée de Fourier numérique précise pourrait être d'une grande aide, par exemple, dans l'étude des dérivées fractionnaires. On peut envisager aussi la généralisation à  $n$  dimensions des méthodes mathématiques évoquées dans cette thèse, ce qui, par exemple, pour  $n=2$ , semble déjà annoncer des perspectives intéressantes dans le traitement d'image. Ou bien imaginons, pour l'analyse par ondelettes (wavelets), un affinement semblable à celui présenté dans le premier article pour les transformées de Fourier. Enfin, le fait d'avoir facilement accès aux polynômes splines d'ordre élevé, avec leurs conditions frontières idéales, apportera un outil puissant à l'analyse numérique ou au traitement d'image. Et dire qu'on n'a même pas encore injecté les logiques floues dans tout cela! Bref, les idées sont nombreuses, les rêves sont au rendez-vous. Il ne reste qu'à les matérialiser!

## BIBLIOGRAPHIE GÉNÉRALE

1. E. O. BRIGHAM. The Fast Fourier Transform. Prentice-Hall. Englewood Cliffs. 1974. pp.11, 98.
2. J. SCHUTTE. Rev. Sci. Instrum. **52(3)**, 400 (1981).
3. S. MAKINEN, Rev. Sci. Instrum. **53(5)**, 627 (1982).
4. S. SORELLA and S.K. GHOSH. Rev. Sci. Instrum. **55(8)**, 1348 (1984).
5. M. FROEYEN and L. HELLEMANS. Rev. Sci. Instrum. **56(12)**, 2325 (1985).
6. J. N. LYNESS. Math. Comp. **24**, 101 (1970).
7. J. N. LYNESS. Math. Comp. **25(113)**, 59 (1971).
8. J. N. LYNESS. Math. Comp. **28(125)**, 81 (1974).
9. J. N. LYNESS. J. Comput. Phys. **54(1)**, 57 (1984).
10. G. GIUNTA and A. MURLI. ACM Transactions on Mathematical Software, **13(1)**, 97 (1987).
11. J. N. LYNESS. QUG1: Trigonometric Fourier Coefficients. TM 370, Argonne National Laboratory. Argonne. 1981.
12. W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, B. P. FLANNERY. Numerical Recipes in FORTRAN. (2nd Ed.) Cambridge. New York. 1992. p.577.
13. N. MORRISON. Introduction to Fourier Analysis. Wiley-Interscience. New York. 1994. Chap. 13.
14. E. SCHIEFERDECKER and K. STREHLKE. *In* Fundamentals of Mathematics. Vol. 3. Analysis. H. BEHNKE, F. BACHMANN, K. FLADT, W. SÜSS, S.H. GOULD. The MIT Press. Cambridge. 1986. p. 80.
15. E. BUTKOV. Mathematical Physics. Addison-Wesley. Reading. 1968. p. 226.
16. E. KREYSZIG. Advanced Engineering Mathematics. (Fifth Ed.) Wiley. New York. 1983. p. A-55
17. J. H. AHLBERG, E. N. NILSON, J. L. WALSH. The Theory of Splines and Their Applications. Academic Press. New York. 1967. p. 109.

18. C. F. GERALD, P. O. WHEATLEY. *Applied Numerical Analysis*. (Fourth Ed.) Addison-Wesley. Reading. 1989. p. 211.
19. J.E. RUDZKI, J.L. GOODMAN and K.S. PETERS. *J. Am. Chem. Soc.* **107**, 7849 (1985).
20. D.A. HUTCHINS. *Can. J. Phys.* **64**, 1247 (1986).
21. L.A. MELTON, T. NI and Q. LU. *Rev. Sci. Instrum.* **60**(10), 3217 (1989).
22. S.E. BRASLAVSKY and K. HEIHOFF. *In Handbook of Organic Photochemistry*. J.C. SCAIANO. CRC Press. 1989. p. 327.
23. K.S. PETERS, T. WATSON and K. MARR. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 343 (1991).
24. S.E. BRASLAVSKY and G.E. HEIBEL. *Chem. Rev.* **92**(6), 1381 (1992).
25. P.R. CRIPPA, A. VECLI and C. VIAPPIANI. *J. Photochem. Photobiol.* **24**, 3 (1994).
26. K. MARR and K.S. PETERS. *Biochemistry.* **30**(5), 1254 (1991).
27. M. ROHR, W. GÄRTNER, G. SCHWEITZER, A.R. HOLZWARTH and S.E. BRASLAVSKY. *J. Phys. Chem.* **96**(14), 6055 (1992).
28. D. BRUCE and O. SALEHIAN. *Biochimica et Biophysica acta.* **1100**, 242 (1992).
29. K.S. PETERS, T. WATSON and T. LOGAN. *J. Am. Chem. Soc.* **114**(11), 4276 (1992).
30. P.J. SCHULENBERG, M. ROHR, W. GÄRTNER and S.E. BRASLAVSKY. *Biophysical Journal.* **66**, 838 (1994).
31. J-L. HABIB JIWAN, A. K. CHIBISOV and S.E. BRASLAVSKY. *J. Phys. Chem.* **99**, 10246 (1995).
32. M.E. van BREDERODE, T. GENSCH, W.D. HOFF, K. J. HELLINGWERF and S.E. BRASLAVSKY. *Biophysical Journal.* **68**, 1101 (1995).
33. D. ZHANG and D. MAUZERALL. *Biophysical Journal.* **71**, 381 (1996).
34. C.K.N. PATEL and A.C. TAM. *Reviews of Modern Physics.* **53**(3), 517 (1981).
35. A.C. TAM. *Reviews of Modern Physics.* **58**(2), 381 (1986).
36. J. SCHUTTE. *Rev. Sci. Instrum.* **52**(3), 400 (1981).
37. S. MAKINEN, *Rev. Sci. Instrum.* **53**(5), 627 (1982).



38. S. SORELLA and S.K. GHOSH. *Rev. Sci. Instrum.* **55**(8), 1348 (1984).
39. M. FROEYEN and L. HELLEMANS. *Rev. Sci. Instrum.* **56**(12), 2325 (1985).
40. N. BEAUDOIN. *Canadian Journal of Physics.* **76**(9), 659 (1998).
41. J. C. ANDRE, L. M. VINCENT, D. O'CONNOR and W. R. WARE. *J. of Phys. Chem.* **83**(17), 2285 (1979).
42. J. N. DEMAS. *Excited States Lifetime Measurements.* Academic Press. 1983. p. 148.
43. P. J. ROUSSEUW. *Handbook of Statistical Methods for Engineers and Scientists.* Edited by H. M. WADSWORTH. McGraw-Hill. 1990. p. 16.1.
44. S. E. BRASLAVSKY. *The Spectrum.* **7**(4),10 (1994).
45. P.J. SCHULENBERG, W. GÄRTNER and S.E. BRASLAVSKY. *J. Phys. Chem.* **99**(23), 9617 (1995).

— Jusqu'à la présente ligne, cette bibliographie-ci reproduit exactement celles à la fin des articles —

46. B. HANKAMER, J. BARBER, and E.J. BOEKEMA. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **48**, 641-671, (1997).
47. L. K. LANYI. *Biochim. Biophys. Acta.* **1183**, 241-261, (1993).
48. C. BRAÜCHLE, N. HAMPP and D. OESTERHELT. *Adv. Materials*, **3**, 420-428, (1993).
49. R. R. BIRGE and T. M. COOPER. *Biophys. J.*, **42**, 61-69, (1983).
50. A. WARSHEL and N. BARBOY. *J. Am. Chem. Soc.*, **104**, 1469-1476, (1982).
51. R. E. CLAPP. Internal report # MTR10230, The Mitre Corp., Bedford Mass.
52. H. J. SASS, R. GESSENICH, M. H. J. KOCH, D. OESTERHELT, N. A. DENCHER and G. BULDT. *Biophys. J.*, **75**, 399-405, (1998).
53. F. BOUCHER and R. LEBLANC. *Photochemistry and Photobiology.* **41**(4), 459 (1985).
54. H. GARTY and S. R. CAPLAN. *Biophys. J.* **37**, 405 (1982).

## ACCUSÉS DE RÉCEPTION

### Premier article.<sup>19</sup>

From razavy@bethe.phys.ualberta.ca Wed May 28 10:04:19 1997

Received: from triton.uqtr.quebec.ca by neptune.UQTR.UQuebec.ca (AIX 4.1/UCB 5.64/4.03) id AA71908; Wed, 28 May 1997 10:04:18 -0400

Received: from relay.Phys.UAlberta.CA by triton.uqtr.quebec.ca (8.8.4/8.8.4) with ESMT-Pid KAA38706; Wed, 28 May 1997 10:04:17 -0400

Received: from bethe.phys.ualberta.ca ([129.128.7.141]) by relay.Phys.UAlberta.CA with SMTP id <57987-2>; Wed, 28 May 1997 08:01:54 -0600

Received: by bethe.phys.ualberta.ca (NX5.67c/NeXT-2.0) id AA05846; Wed, 28 May 97 08:04:15 -0600

Date: Wed, 28 May 97 08:04:15 -0600

From: razavy@phys.ualberta.ca (Mohsen Razavy)

Message-Id: <9705281404.AA05846@bethe.phys.ualberta.ca>

Received: by NeXT.Mailer (1.87.1)

Received: by NeXT Mailer (1.87.1)

To: Normand\_Beaudoin@UQTR.UQuebec.ca

Subject: QM 116

Status: R

Dear Dr. Beaudoin

Re: File # QM 116

This will acknowledge with many thanks the receipt of your paper "A high accuracy mathematical and numerical method for Fourier transform..." submitted for publication in the Canadian Journal of Physics.

The submission of any manuscript to Can.J.Phys. constitutes the authors' guarantee that no part of the paper reporting original work is being considered for publication elsewhere under the names of any of the authors of this manuscript and that none of the authors will submit such material to another publisher unless it is rejected or formally withdrawn. The submission of a manuscript is also taken as the authors' assurance that, to the best of their knowledge, none of the contents of the paper represented as reporting original work by the authors has been previously published or submitted for publication by others. Your paper is currently under review and we shall inform you as soon as possible concerning the referees' recommendations. Should your paper be accepted, if you are able to provide a copy of the final version of your manuscript on a diskette labelled with the file number and software used it will speed up the publication process. I also include a copy of our Instruction to Authors noting changes which must be made before publication if your paper is accepted.

Yours sincerely,

M. Razavy,

Associate Editor

---

19. Depuis, cet article a été publié. Voir [40].

## **Deuxième article.**

From DHOGAN@biophysics.faseb.org Wed Sep 23 14:07:35 1998  
Received: from triton.uqtr.quebec.ca by neptune.uqtr.quebec.ca (8.8.4/8.8.4) with ES-MTP id OAA45352; Wed, 23 Sep 1998 14:07:35 -0400  
Received: from faseb.org by triton.uqtr.quebec.ca (8.8.4/8.8.4) with ESMTTP id OAA42820; Wed, 23 Sep 1998 14:07:34 -0400  
Received: from biophysics.faseb.org (biophysics.faseb.org [192.0.2.83]) by faseb.org (8.8.8/8.8.8) with SMTP id OAA11598 for <Normand\_Beaudoin@UQTR.UQuebec.ca>; Wed, 23 Sep 1998 14:00:24 -0400 (EDT)  
Received: from DOMAIN-Message\_Server by biophysics.faseb.org with Novell\_GroupWise; Wed, 23 Sep 1998 13:59:06 -0400  
Message-Id: <s608feaa.020@biophysics.faseb.org>  
X-Mailer: Novell GroupWise 5.2  
Date: Wed, 23 Sep 1998 13:58:52 -0400  
From: "Debbi Hogan" <DHOGAN@biophysics.faseb.org>  
To: Normand Beaudoin <Normand\_Beaudoin@uqtr.quebec.ca>  
Subject: B80749  
Mime-Version: 1.0  
Content-Type: text/plain; charset=US-ASCII  
Content-Disposition: inline  
Content-Transfer-Encoding: 8bit  
X-MIME-Autoconverted: from quoted-printable to 8bit by neptune.uqtr.quebec.ca id OAA45352  
Status: RO

Dear Dr. Beaudoin:

Your manuscript was received in our office on 8/28/98 and has been assigned the manuscript number B80749. I am sorry that you have not received an acknowledgement letter - our records indicate that one was mailed to you.

We have received one of the reviewer's comments on your manuscript and hope to receive the second review within a few days. As soon as we have a decision, we will fax it to you. Our records do not list a fax number for you - please provide a number so that we can send the reviews to you as quickly as possible.

Cordially,  
Debbi Hogan

BIOPHYSICAL JOURNAL EDITORIAL OFFICE  
9650 Rockville Pike  
Bethesda, MD 20814  
phone: (301) 571-0663  
fax: (301) 571-0667  
e-mail: dhogan@biophysics.faseb.org

Le vendredi 28 mai 1999

Nous, les soussignés, membres du jury de la soutenance de la thèse de doctorat en biophysique de monsieur Normand Beaudoin, déclarons que ce dernier a soutenu devant nous, avec succès, sa thèse de doctorat intitulée:

**Méthode mathématique et numérique de haute précision pour le calcul des transformées de Fourier, intégrales, dérivées et polynômes splines de tout ordre. Déconvolution par transformée de Fourier et spectroscopie photoacoustique à résolution temporelle.**

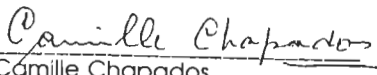
REMARQUES:

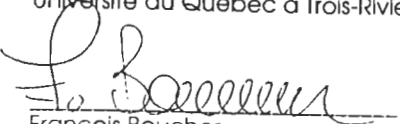
Nous, membres du jury, reconnaissons que les méthodes mathématiques développées par le candidat sont originales et doivent être largement publiées après en avoir revu le développement formel. Nous reconnaissons également la maturité mathématique du candidat ainsi que son imagination fertile qui le place aux premières lignes d'un secteur exigeant.

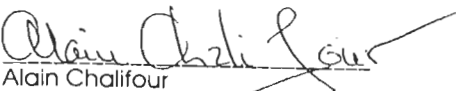
Nous soulignons la persévérance et l'acharnement du candidat dans la poursuite de l'objectif de son projet de recherche.

Nous recommandons que le travail héroïque du candidat dans le domaine de la biophysique qui a conduit à des résultats prometteurs soit poursuivi.

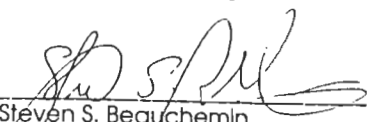
Pour toutes ces raisons nous décernons au candidat la mention *excellent*.

  
Camille Chapados  
Président du jury  
Université du Québec à Trois-Rivières

  
François Boucher  
Directeur de recherche  
Université du Québec à Trois-Rivières

  
Alain Chalifour  
Université du Québec à Trois-Rivières

  
Serge Gauvin  
Université de Moncton

  
Steven S. Beauchemin  
University of Pennsylvania